Structure of everninomicin (Ziracin)

Figure 1

Overlapping cosmids and clones spanning 185 kb of chromosomal DNA containing the Everninomicin Pathway region.

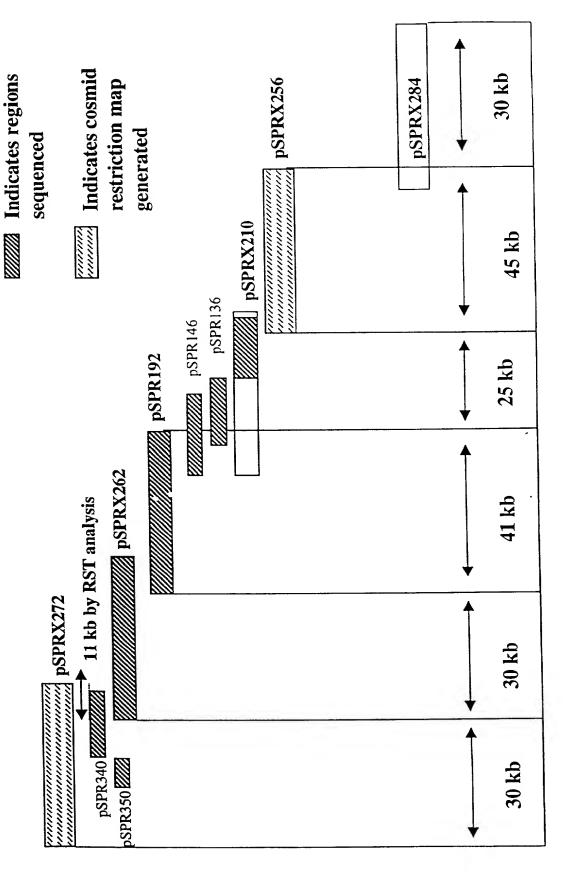
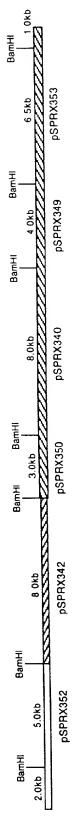


Figure 2A

pSPRX272 37.5kb Cosmid pSPRX272

Regions sequenced indicated by crosshatches.

Fragments cloned indicated by clone designation benigth fragment.



pSPR272 (37500 bp)

Figure 2B

Cosmid pSPRX256

Regions sequenced indicated by crosshatches.

Fragments cloned indicated by clone designation benigth fragment.

Cosmid pSPRX256

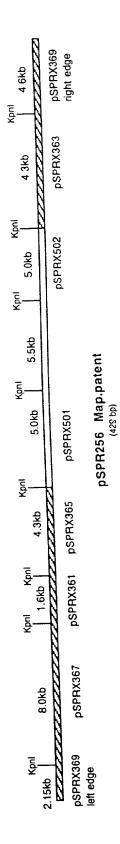


Figure 2C

Figure 3 (A)

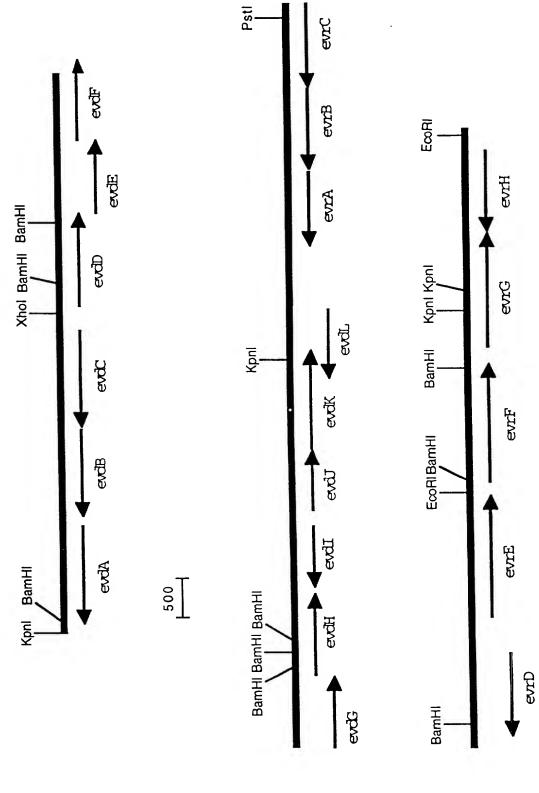


Figure 3 (B)

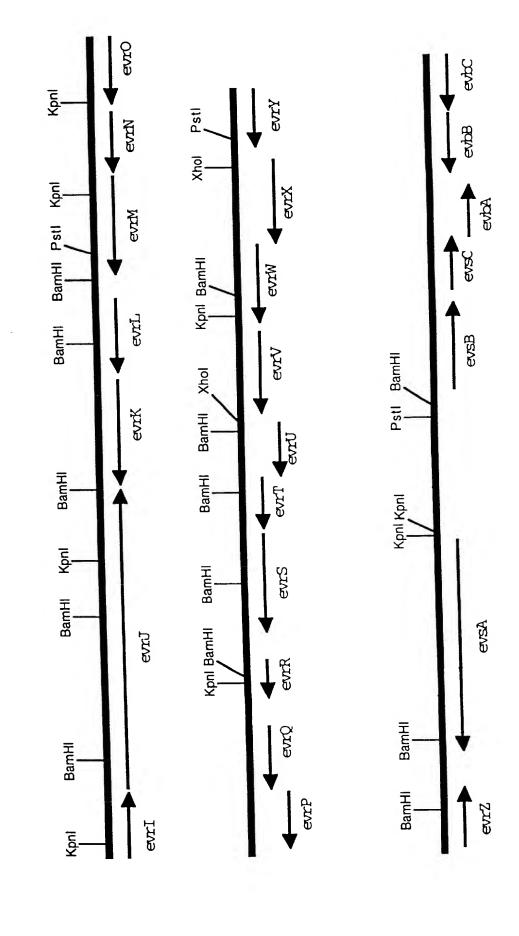


Figure 3 (C)

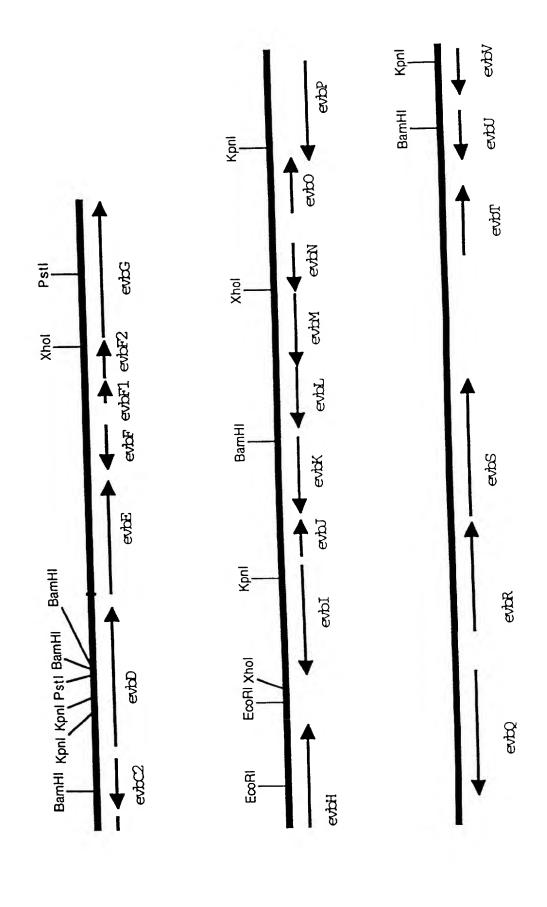


Figure 3 (D)

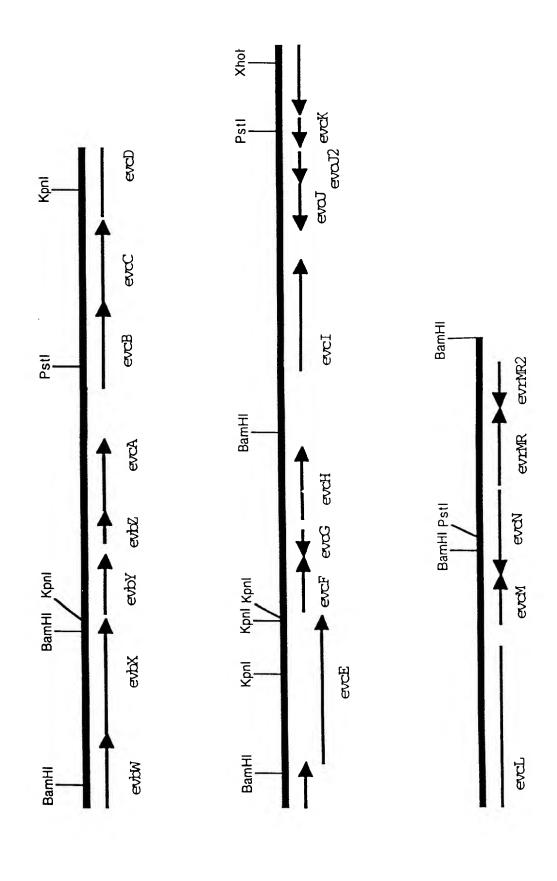
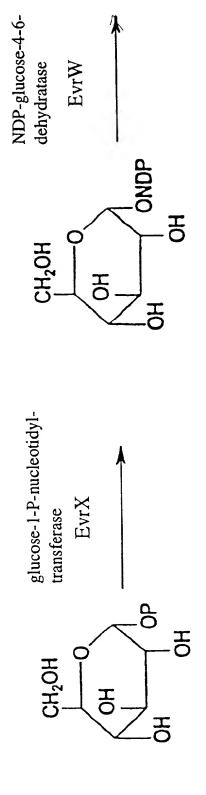


Figure 4(A)

Figure 4 (B)

Figure 5 (A)



NDP-D-glucose

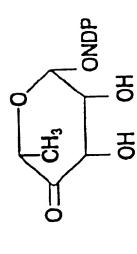
D-glucose-1-phosphate

Figure 5 (B)

EvrV

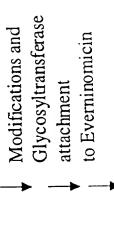
ONDP

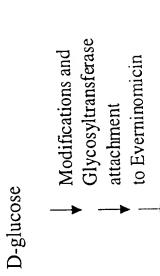
F



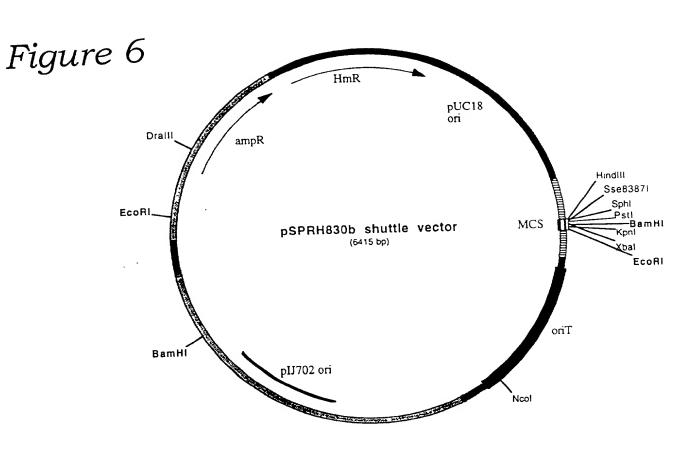
NDP-4-keto-rhamnose

NDP-4-keto-6-deoxy-





pSPRH830b E.coli-Micromonospora shuttle vector

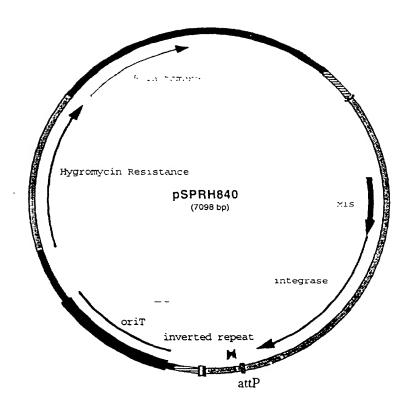


pSPRH830b - pSPRH826b backbone

Function	Source
- Ampicillin resistance	(pUC18)
- Multiple cloning site	(pUC18)
- pUC18 origin	(pUC18)
- Hygromycin resistance	(p16R1)
- oriT (origin of transfer)	(pRL1058)
- pIJ702 origin of replication	(рИ702)

pSPRH840 integrating vector

Figure 7A



pSPRH840 - pSPRH826b backbone, pMLP1 xis, int attP insert

pSPRH840 conjugated from E. coli into	HmR transformants obtained
M.carbonacea	+
M.rosaria	-
M.halophitica	+

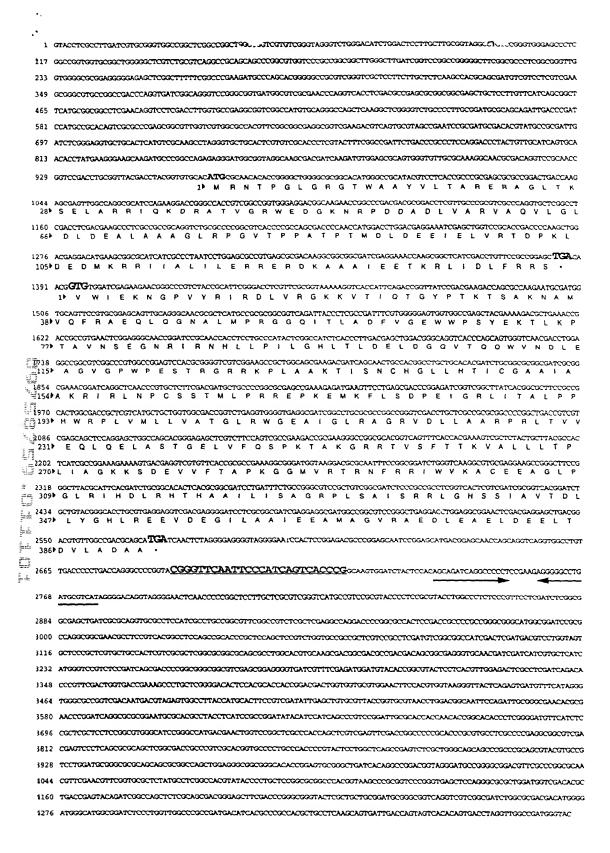


Figure 7B

pSPRH826b Insertion plasmid

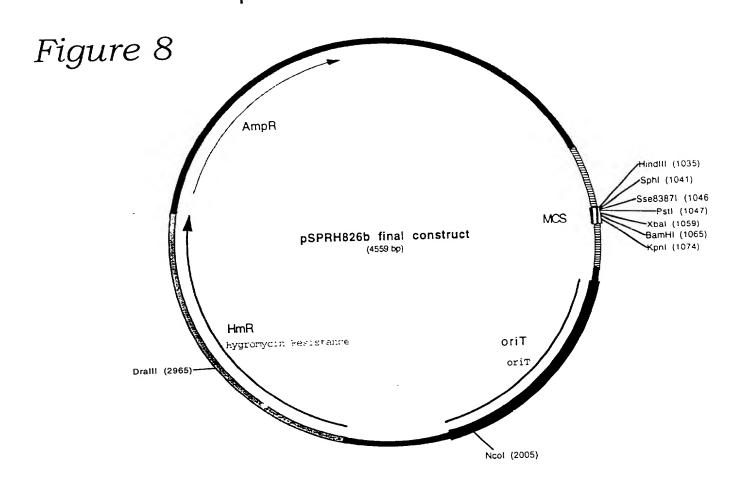


Figure 9

Analysis of M. carbonacea and M. halophytica pSPRH840 insertion site AttB/AttP region

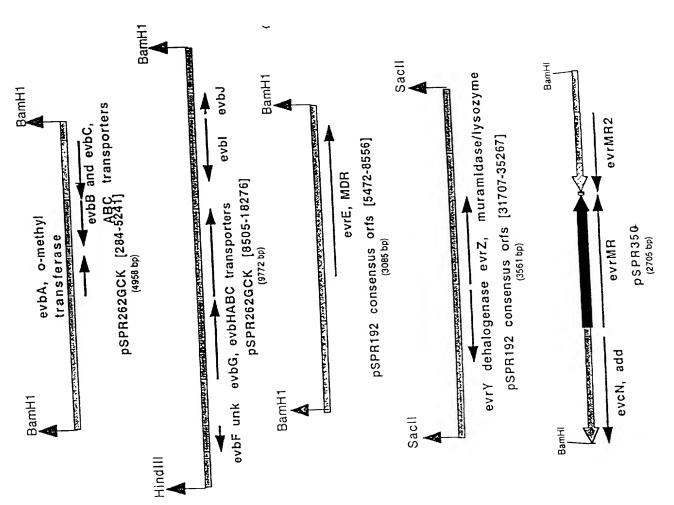
```
8 8 8 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            179
175
175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cogniticality control constant constant constant constant constant control co
                                                                                                                                                                                                                                                                                                                                                        TGATCAACTCTIAGGGAAGGGGAAATCCACTCCGGAGACGCCCGGAGCCAATCCGGA
TGATCAACTCTIAGGGAAGGGGAAAT-CACTCCGGAAGGCCCCGGAACAATCCGGA
TGATCAACTCTIAGGGAAGGGGAAATCCACTCCGGAAGCCCCCGGAGCCAATCCGGA
TGATCAACTCTIAGGGAAGGGGAAATCCACTCCGGAGACCCCCGGAGCAATCCGGA
                                                                                                                                                                                                                                                          TATCAACTCTAGGGAAGGGGAAATCCAACTCCGGAAGGCCCGGAAGCAATCCGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CENTENCIOLACENCENDETENGITOCOCTIOTITANCOCOCTIONCENDEDEDEDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cccritcaattcccatcaccccdacdiaasaashergerxchvskccrswkcosabsrg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Insertion juncture
Alignment of pMLP1 attP region with religation clone edge sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CCTTCCCCT-TCCTCACCTTCCC-\
CCTTCCCCCT-TCCTCACCTTCCC--\
CCCCCTCCC----MGACCCCCTCAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CCCYYCKVCGYATYHWSAOCGICKCSYGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        coct-caccattttcaccc----
                                                                                                                                                                                                                                                                                                                                                                                                H. halophytica Pati relig-9
H. carb Psti relig-1
H. carb Psti relig-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 H halophytica Psti relig-9
H carb Psti relig-1
H carb Psti relig-4
                                                                                                                    M.halophytica PstI relig-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        halophytica PstI relig-9
                                                                                                                                                                                                                                               philp introy att region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pacel inttch.att region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMEPT INCTCA. ALL region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                prof. intTCA, att region
                                                                                                                                                               M.carb Pstl relig-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                M.carb PstI relig-1
M.carb PstI relig-4
                                                                                                                                                                                                       M.carb PstI relig-4
                                                                                                                                                                                                                                                                                                                        Consensus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Consensus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Consensus
```

3 . . NMLP1 attP region

1 -natical contractions and considered considered considerations and λ_1 and λ_2 and λ_3 and λ_4 and λ_3 and λ_4 and λ_4

¹⁷⁴ AGGCCCCTCCGAAGAGGGGCCTGATGCGTCATAGGGGACAGGTAGGGGAACTCAA





1 GGTACCCGACCGTGTCCCGGAACAACGAGTCGAGATACGGCGAGAGGAACACCCCCCGGTAGTCCGGGTAGACGGTGGGCGCGAAGGCGTAC 93 GCGCCTTCGACGGTCAGCGGGCGGACACCGGCGGGTCAGCTCGTCACGTGTACGCGGGGGACGTACAGGATCCACTGTCCGCCAGCC

< * S T V H V R P V Y L I W Q G G A 460 AGCAGGTCGGTGCGCATGCCCTCGACGTCGGCGGCGAACCTGTCGAGCCGGCTGCGGTCGGCGACCCCTCGGGCGTCCTCCTCGCCGATCAG RDLRSRDAVGRADEE 552 CGCGGCCACCGGCTCGGCCGGCTGCCGCGCGCGCACGGGCGATGGTGTAGCGGACCTCCCCACCGTGCACCGGAAGGCGTTCCACGTCGA < A A V R E A P Q R A G A R A I T Y R V E G G H V P L R E < L A F G F R A A L A Q V S R A S F F F F H E D Y I Q D F S T
736 TTGTCCAGGATGTCCCCGAGGTACGGGTCCTCGAAGACGCCCGTCCGGTGCCAGCAGCGCGTCAACGCCCCGGAGGATGGAGTCGAG</pre> 1012 ACGGTACGCAGCATCACCCCGTCGTTGCAGCCGATCTCCACCACGAACGGGTCCGGGCCGGTGGCCTGCTGCTCCAGCAGGTGCCGCGGGT

V T R L M V G D N C G I E V V F P D P G T A E H E L L H R A T 1104 GTCGGCGAAGTGCTTCTGCATCACGGAGGAGCCCGACGAGTGGTACGGGTAGTCCTGGTGGAACATCTCCTCCCGGGGCACCTCCTCCATGA < D A F H K Q M V S S G S S H Y P Y D Q H F M E E R P V E E M L
1196 GCTGCACCATCGTGCAGCCCGCGAGAGCCCGCCACCGCCAGGTGGAAGAAGAACTCGTCCGCGAGCTGGTCCTCGGTGAGGAACCGGTCGGAG</pre> 1560 TACATGGGAGGAGGAGAAGATCTCGCCGGCCAACCGTTCCGTGGTGGGGAGCCGGTCTGGTAGCCGAAGCTTGGCGAAGCCGGTCATGGT
<Y M P L S F I E G A L R E T T P L S G T Q Y G L K A F G T M T < H V P W P Y S I N L A I D R A A L O E L I R D R A P H R V</p> 1744 AGACGTACCAGACGTGTTCGTTGTCGGTCGCCGTCACCGGCAGGTCAGGCCGACCTGGTCGGCGATGTCGGCCAGGCCCTCCTCGTAGCGC V Y W V H E N D T A T V P L T L G V Q D A I D A L G E E Y 1836 CGGGCCACGGCGGCGCGCGGCGATGTAGTCGTCGAGCCGGCACAGCTTGCGGCGCAGGATCTCGGCCTGCACCTCGTCGAGGCGGCAGTT 1928 GTGCCCGGCGTCTCGACGACGTAGTAGACCTGTTCCATGCCGTAGTAGCGCAGCCGCCGCAGCCGCTCGTCGATCACGGCGTCCGCGGTGA
 < H G P T E V V Y Y V Q E M G Y Y R L R R L R E D I V A D A T V 2020 CCACGCCGCCGCCGCCGTACGCGCCAGCACCTTCGTCGGATAGAACGAGAAGGCAGCGGCGTCGCCCATCGTGCCGGCCAGCCGTCCG $\,$ < V G G G D G Y A G L V K T P Y F S F A A A D G M T G A L R G 2112 CCCCGACGCGCCGCTGCGACTCCTCCAGGACCTTGAGGCCGTGTTCCCGGGCCACCCGCAGCACCGGGTCCATGTCGACGCA 2296 GGACGTCCACGAAGACCGGCGTCGCGCCGGCCGGTCGATGGCGACGCGCGGCGGCGGCGTGTTGGAGACGGTGATGACCTCGTCGCCCCCC V D V F V P T A G V A D I A L V T P A A T N S V T I V E D G 2388 GGCCCGACGTCGAGCGCCTGGAGTGCGAGCTTGATGGCGTTGGTGCCGTTGTCCACCGTGACGCAGTGCGGCATGTCGTGATAGGCGGCGAA <P G V D L A Q L A L K I A N T G N D V T V C H P M D H Y A A F 2480 CTCCTGCTCGAAGCCGCCGCACGCCTCGCGCCGCGACGACGACTCGAAGACCGTCTGCACGGCGTCGAGGAGGTCGTCCCGTTCCT < D A N T L A A G V T T R L A Y L R T D L G A V L A R A G A R</p> 3029 GCTCGCCGCGCGCCCCGCGCAGAAACCGACCGCGATGTCCCGGGCCGCCTGGGCGATGCCGGCGTAGATGCCGAGCATGGTGATCGAGCTG

< A A E A S I R P I V G A L L D V V T R G A P T L P A R L D A 3857 CCACCCATCTCCTCGGTCAGATTAGACATCGCCTGCTTCCGTTCGCGCTGTGCCGAACCTGTCGCTATCAGGGTGCGCGGCGATCACC 3947 AATTGCTGGCTGATTGTCCACCGACGATGCTCGACAGGGGATACCCAGAATAGGCGGCAACGCTTGGCGAAACCCCTGTCGTGCGCCGGAGT 4039 TCGCCAAGTCTAGTTGGATCACTCGAGCTTCCCCCGCAGAAGCCGTGAACCATGGGCCAGCGGTTGACGTCGCTATATATTCGGCCGACACG RTRLADGIAKVYKWVEADEPVRGER• > M S R P R I L V A G N F H W Q A G F S Q I V A A G N F H W Q A G F S Q I V A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A A G F S Q I V A G T

7805 CTGGTGGCGGTGGGAGCCGATGACCGAGATGGTCGAGGCCGCGGCGGCGGCCGCCGCTGCGCCGCTGCGGTGCGGACGCTGGT W W R W E P M T E M V E A A A A R P P L R R L R V C G R W >W D G G S C A G F E E A T L S E P G W L R A R G V E V H P P 7989 CCGTTCGGCCACGTCGACCACATGGGCCGGTCGCTGATCTCACCGGTCCTGGTGCGGCCGCTGGTCACCAGCACCAGCCCTGTTGACCCC PFGHVVEQMGRSLISPVLVRPLVTSTG 8081 CCGGATGTTCGAGACGCTGGCCTCGGGCAGCCTGCCCGTGCTCCCGGTCGCCGCAAGTTCCTCGCGCCGGTCTACGGCGACGAGGCGGAAC > R M F E T L A S G S L P V L P V A A K F L A P V Y G D E A 8173 ACCTGATGCTCGGCGACGACCCGGCCGGAACGCTGAGCCGGCTCTCGGCCGAGCACGAACGGTACGGACGACTGGTCGGTGAGATTCAGGAC >H L M L G D D P A G T L S R L S A E H E R Y G R L V G E I Q D 8265 CGGCTCCGCGTCGAGCATCACCCTCGCGGGACCTGCTCGATCTGCTGGCCTGAGGAATGAGGAGCAGATGACCCCCCTG > R L R V E Y G Y P R V L R D L L D L L A 8354 CGGATCGCGATGGTCAACATACCGTTCCGGTTGCCGAGCGACGAGCGGCAGTGGATCACGGTCCCGCCGCAGGGGTACGGCGGGATCCAGTG

> R I A M V N I P F R L P S D E R Q W I T V P P Q G Y G G I Q W 8446 GATCGTGGCCAACAAGATCAAGGGCCTGCTCGAACTCGGGCACGAGGTGTTCCTGCTCGGTGCCCCGGGCAGTCCGCGTACGCATCCACGCC I V A N K I K G L L E L G H E V F L L G A P G S P R T 8538 TGACCGTGGTGCCGGCGGGCGAGCCCGAGGACATCCGGGCATGGTTGAAGTCCGCTCCGGTGGACGTCGTCAACGACTACAGCTGCGGCAAG > V D P I E L P P G V G L V A S H H M T T R P S Y P A G C 8722 CTCGAAGGCGCAGCGGAGCAGTGCGGCGGCGGGGGGGGGCGCCCGGTCATCCCGATCGGGGTGGATCCGTCGCTCACCGCCCCGGCGACC >R K D D F L L F M G R I S P F K G A L E A A A F A R A A G R R 8906 CTACTGATGGCGGTCCGGCCTGGGGGGGGGTACCTCGACCGGATCATGGGCGAGTACGGCGAGCTACGCTCACCCTCGTCGGCGAGGTGGG > L L M A G P A W E P E Y L D R I M G E Y G D H V T L V G E V G 8998 GGGTCAGGAACGTATGGACCTGCTCGCCACGGGGGCGGCTGCCATCCTGGTGCTCTCCCAGCCGGTGCCCGGCCCGTGGGGCGCACGTGGTGCG > G Q E R M D L L A T A A A I L V L S Q P V P G P W G G T W C 9090 AGCCGGGTGCGACGACGAGCAACGGCTGCCTGGCGAGAACGGCTGCCTGGCGAGAACGGCTGCCTGGCGAGAACGGCTGCCTGGCGAGAACGGCTGCCTGGCGGAGATCGTGCCGGCC >E P G A T V V S E A A A S G T P V V G T S N G C L A E I V P A > A A I R C W G H V E I A R R Y E A V Y R D V L A G A R W S • 9365 GCCGGCCGGGCGGGCGGCTACGGTGGGGACCGTAGGGGGGTGCCCGCCGCACGCGGAAGCGCCGGTGTCGGCGGTCCGACACCGGCGCCCCC <Y A A R L W A V E A D D A L Q P L T V F G D R R Y A E L R R I
10192 GACGACCTCCGCCACCAACAGTCCCCACGGCCATTTGACCACCTCTCGGAATAGCCTGTCCGCGAATAAACCATACGGTAGGAACAGCGCG
< V V E A G V T G V A M</pre> 10282 GCGATACCGCTCCCGAGCGGGAAATAGGGATTCGACTAGTATTCGGTCCGCGCCGCTGCCAGAACGGCACGCGCTCTCGATTGTCCATTCAT 10374 CCCCGTGCGAGACTCGCCTCGATGTCCTCGATGTCGGTGGGGGGTTTGGGATGACCGGGCACAGCGCCGTCGCGCTGGACGTCGGCGGGGGT T D H D V L P A L G L G C P V A F V L P D P A Y R R P 11017 TCCGGACCTGGTCCGGGTCTACACGGAGCTGAGGGCGTTCCCGCACCGGCTCCCCGCGCGGACGCCCGGGTCACCACCGTGGCGTCCCTGG P D L V R V Y T E L R A F R T G S P P A D A R V T T V A 11109 CGGCCCTGGCCGACTCTCCCCTGACGAGTGCCACCCCGCGTTCGAACGCCGGCACCGGCGGACTTTGACGAAGGAGTGCAGTTGCGACGCC 11291 TGCCGCCCACCCGTCGACGGCCGGCCCGCGCCGGGGCTCCGCGAGGCGCAGGTCGTCTACTTCCTCGCCGCCCGGCTGAGCCCGGCGC

```
11475 TTCGTCCTGGCCAGCTCGGCGGGGGGGGGGGTGTACACGCCGACGGTGTGGCCGCCTACCACGAGGGTCGGCCACCGGGCCCGCCTCGGCGTA
11751 GCCGTGGTGCGGGACTACGTACACGTCGACGACGTCACCGCGATCATGGAGGTCATCGCGCAGCGGCCGGTGACGGCGACCGGGACCGGCT
11843 GCCCACGGTCGTGAACGTCGGGCCTGGCCCACCTCCCTCGCCGAGTTGCTCCAGACGATGTCCACGGTGGCCGGTCGTGAGCTGGAGC
11935 TCATCCGGGACGTCCGCCGGCAGTTCGACCATCGGGGCAACTGGCTCGACACCACCTCGCCCGGGAGACCCTGGGCTGGCAGGCGCGGATC
> S L P D G V R Q C W E A V L T R A G G P G G S P A R P S A R
12118 TCGGGAGAGCGTCTCGGGGGCGGGAACCGCCGCAACCGCGCCCTTCGCAGCAGTTCGTGGCTCAACCCGGCGGCGGTCGCCGCGGTGTAGCC
   >L G R A S R G R E P P Q P R P S Q Q F V A Q P G G G R R G V <E P S R R P P P F R R L R A R R L L E H S L G A A T A A T Y
                                                          junction marker
> E G Q W Q G E P G F R G R T V P V P F L A Q A G P C L V < L A L P L A L R S E P A T R H R D W E K R L G A R A Q N
12302 CCTCGCACCGCCCTGCCAGTACGCCCGCCGCAGCAGCTACCGGGGGGTCAGCCGGCCCGGGTCGATGTCATGGGTGACC
   > LAPPLPVRPPQQVPGGQPARVDVMGD
< ECRGQWYARRLLYRPTLRGPDIDHTV
12382 GCGTGGTCCGGGAGCAGTTGCTCGCGGGCGCCGGCCGCCTTCATGGCGCTGATGAAGGAGGTGTCCTCC
   >R V V R E Q L L A G A G G L H G A D E G G V L <A H D P L L Q E R A G A A K M A S I F S T D E
12451 CCTGACTGGAGGTTGCCCCCGGTACGGCTGAGGGCCAGGTCGAAATCCAACCCGTGGGCGTGCGCGAAC
   <G S Q L N G G T R S L A L D F D L G H A H A</pre>
12520 GCGGAGTCCACCCCATGCACGCCCCAGATCTTGATGTTTCCCTGGTCACGGTGCAGCCGACCAGGTGGAACTGGCCGA
<A S D V G M C A G W I K I N G Q D R H W G V L H F Q G S
                      junction marker
A P R A L R T G V V H A G D R L S K R V A D
2 A A A D L R V D D D V F M L H H H P W R A L M A N R S A S L

12787 CCATTGGTGGCACCGAGGATGGCGCATGGTGCCGCGGGGGCCCGGACCTCCTCGGCGACCTCCTCGGCCTCGGCCGTGACGGGCCGGTCCAG

4 C N T A G L I R M T G G A A R V E E A V E E A E A T V P R D L

12879 CAGGACGTAGTACTCGTCGCCGGAGAGCTTGGCGCATGTTGTGCGGGAGGTGTTTCCTGACGTCTCTCCGCCACCGGAACGCGCAGATCCCCACA

4 C N Y Y E D G S L Q A M N H A L H K R V N E V R F A C I A V V

12971 CCATCGGGTGGTCGGACGGATCGCGGCTGACCACAGAGGCGTTGTTCGGCATCGTCCTCGACATGAGCGTTGTGGGGCCGGGGGAGA

4 M P H D S P D R S V V S A N N P M
13245 CGCGATCAGGGCCGGCCGGTTGGGCACTCGCAGGTTCTGCAACAGCTTCCTGACGTGGTACTCGACGCCCTGACGGCTGAGGTAGACCTTGT
13337 TGGCGATATGCACCGTCCGTTCGCCTGCGGCGATGCTTTCGATGATGCGCGCGTCCAGATCGGACAGTGAGAACTTGAGATTCACAACGCCC
13429 CCTTGATTGTGGATCTGTTGGCTCGTACGCGGCCAGACGTCATCGCCCGGACACCCCCCTGAGGTGCCGGTGGACGACGACTCGCTTCCGCT
13521 GGAAGACCGTCAGCGGCGGAAGGCCGGGTGCCCGCCACTCGTCGTCCTGCGCTTGCGGAGCTGTCTCAGTGTGCGGCGATGCCCCCGGCACG
13613 TAGCCCGCCTGACGGGCCTGTCGGCTTGGGCTCCAACTTCGGCAACCCATCGGTTCGCGCTGCACAGAACCAGCGGAGGAACATTGAGTTTC
13797 CTGTCGCTCGCGCTGAACTCACCAATACGCC; AAAGCGTAGCCGGCCCACTGCGGAGCGTCCACCCCCCGAGGATATCGCCAGGCTTCCATG
13889 CAGAACTGGCAGGATCTTTCATCTCAGCCGCACCTGGCGACAAAACCCCTGCTCAAGACCATGAGTAAGCAGGCGGGGAAATCCATGCAGT
13981 GACATGTGTCACTTTAGACAACCCAGCTCCAGCCAGCCCACCACCTGACAAAAGGGGCGGAATCGCGACCAGAGCGACACCAGCACACATTC
14165 CACCGTGAGCAGGCCGACCGTCCGGCAATGCGAACGACGACGTCCACCCCGTCCCGGTAGCCTCACCGCCGCTGGGGGGGCACGGCCCGGA
< * R R S R D
< P P E Q V R E L V S G L H V L R P R C R T R E L A A T V S W T
15084 TCGGGATCGTGACGTCGACGGCAGCTCGGCGTCGGCGTCGCGTCGCGTGGCGTTGACCACCGCGTCGGGCTGCTCGGCGTCG</pre>
```

```
15725 GAAGGCGATCTGTGCGGTCCGCCCGTCCGGCGTACACAACAGCGCGGCACCGAAACGTCCACGCGGGATCGGGGTCCTCCCGGAGGGTGG

< F A I Q A T R G D P T C L L A A G S V D V G R D P D E R L T A
15817 CCGCCACCACCTCCGGCTCCTCCGGCAGGAAAACCGGGCCCCAACGGGTAGACGCCCAGATCCAGCAGCGCTCCGCCACCCAGCTCG
                V E P E E P L F F R A A G L P Y V G L D L L A
16001 CCGGCGCACGAAGTGGTGCACCCCATGTCGGAGGAAGGTTGAGGTTGTCCATCAACACGAGCCCACGGGCCTGGGTCAGCACCGCCG
16093 CGGTGTCGACCAGCCGGGTGGTCAGCGGCTTCTCCACCAGCACGTGTTTGCCCGCGGCCAGGGCGCGTTCGATCCAGGTGTGGTGCAGCCCG
        T D V L R T T L P K E V L V H K G A A L A R E I W T H H L G
16185 GTCGGCAGCGGAATGTAGACGGCATCGATGTCCGGGCGGTCGAGAACGGACTGGTAGCCCTCGGCGGCGCGCCCCCGAACTCCGCGGCGAA
YE I N L Y Y G H R L L A E L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D L Q R V T V W C F E D P V E T P F D P V E T P F D W E S R Y F R G G E E T L V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y L V R E P L Q R V T V T D Y 
< A A E L V E A L F L P R P G P Q N D P I C Q V T P G M E M A
16735 TCGAGCAGCCCCGCTGGTAGCGCGCGCGCCACCCCGGTGCCG
<D L L G A Q Y R A H V L L H A V G D I E K V L F A V V G R H R</pre>
16919 GCGCGATCCCGTCCGCGGTGTGCCGCCAGTCCGGCAGCCCTCGTAGCGGCACCGCCGCACGGTCATCTCGTGCCGGCCCTTGGCCCCGGTG
           I G D A T H R W D P L G R L P V R R V T M E H R G K A
<FWSLVSVLDHRGPAGAASRVIAAVAPS
< T E Q A A A Y F A S P L C S L V T R T D M N V L G D V R L L
18298 CCAGCGACACCTTGCGGGTGACGAACTCGGGACCACGCACCGGTGATTCGTGGTTGAACAGAATGCCGGAGACGGCATACATGCCGTACGAC
18390 TCACGGTAGTTCTGCACCATGTAATGCCCGAACGCCTTGGCCGCCGTACGGGCTACGGGGTGGAACGGGGTCAGCTCATTCTGGACGG
<ERYNQVMYHGFAKAAAGCGCGGCTGACCGGCTGCGGGACTCCGGCACCTCCGACGCCCCCA
19217 ACCGTTTTGTCCCCCTAACGTCGGCGAGGCTGCCAGCCGGCCCGCGAGCCGGGCTTCACGAAGGCGCCAACTTCCGGTGAGAGAGCAG
19309 GGCTCATGTCGCAGAGCCGGCGGCGGCAGCACCCCCAGCCCGCCTGCCACGTCGCCGCCATGGTCGCGTTGGTCGCGGTGATGAT
19400 CCCGATGGTGCTGGCCACCCTCGACAACACCATCATCGGCACCGCACTGCCCACCGTGGTCGGCGAGTTGGGCGGCCTCAGCACGCTCTCCT
> P M V L A T L D N T I I G T A L P T V V G E L G G L S T L S

19492 GGGTGATCACCTCGTACACGCTGGCCACGGCCGCCTCCACGCCGGTCTGGGGCAAGCTCGCCGACATGTACGGCGGCAAGGTGGTCTTCGTG
>W V I T S Y T L A T A A S T P V W G K L A D M Y G G K V V F V

19584 GCCACGCTGGTCGTGCTGCCGGGTCGCTGTCCGGCATGGCGCAGGCATCACCCAGCTGACCGTCTTCCGGGCCGTGCACGGGCT
```

> A T L V V F L A G S L L S G M A Q S I T Q L T V F R A V H G L 19676 CGGCGGGGGCTGATGGTCTGCGGGTTCGCCATCATGGTGGAGGTTCTCGCGGCCCTGACCTGCCCAAGTACCAGGGCATCATGTCGG

S G A G G L M V C A F A I M V E V L A G P D L P K Y Q G I M S 19768 CGACCATGGCCTGACCATGGTGGCGGCCCGCTCGTCGGCGGCCTGATCACCGATGAGCTCGGCTGGCGCTGGTGCTTCTACATCAACCTG
>A T M G L T M V A G P L V G G L I T D E L G W R W C F Y I N L 19860 CCGATCGGGGCGGTCGCGCTCATCGTGGTGCTGATGATGACCCTGCCGCCGACACCAAGGCCCGGATCGATTACGCGGGTGCTGC > PIGAVALLIVVL M M H L P R R H T K A R I D 19952 CCTGCTCACCGTGGTCAGTTCGTGCGTCGTGCTGGTGACCACCTGGGGCGGCATCACCTACCCCTGGGCGTCTCCCGATGATCCTGGGGCTGG L L T V V S S C V V L V T T W G G I T Y P W A S P M I L G 20044 TCGCGCTCGGGGTGCTGACCTGCGCGCTCTTCGTGGTGGTGGTCGAGCGGTGGCCGAGCCGTTGGTGCCCCTGGCCATGTTCCGCAGCCTG >V A L G V L T C A L F V V V E R R V A E P L V P L A M F R S L 20136 AACTTCACCCTGAGCACCTCATCGCCTTCCTGGCCTTCCTGGCCTCATCGCCGTGCACCTTCCTGGCCCTGTTCCAGCAGCAGCGGTGCA G A S A S D S G L L L P L L S M A A V N V V G G R L M 20320 GCGGGCGTTCCTACCGGCTGATGCTCGCCGGTGCGGCGCTGATGACCCTGAGCCTGCTCTCTCGCCCTGATGACGTGGGCACCAGC
SG G R S Y R L L M L A G A A L M T L S L L L F A L M D V G T S 20412 CGGACGGTCACCGCGATCCCCATGGTCGGCCTCGGCGCAGGGCTGGGGCTGATCAGGACCAGCCTGATGGTGGCGCTGAGCAGCGTGGA

R T V T A I P M V G F G A G L G L M Q T S L M V A L S S V E
20504 GATGAGGAACCTCGGGGTGGCCGCCTCCACGTCCACGCTCTTCCGCACCATCGGTGGGGCGGTGGGGGCGTCGGCGACGGTCTCGCTGTTCT > M R N L G V A A S T S T L F R T I G G A V G A S A T V S L F

20596 CCGTGCGGGTGCAGTCGGCGCTGGCCGATCGGGGGGTGGCCGACGTGGCTGACCTCCTCGGCCACTCCGCGGCTGCACGCCGGCTGGACGCCGCCGGGCTG

>S V R V Q S A L A D R G V A D V A D L L G H S A R L D A A G L

20688 GCCCAACTCCCCCGGGCCGTCCGTGTCCACTTCATGCACGCGGTGGCCTCCGGCACCCGGTGGGCCTTCCTGATGACCGTGCTGGCGGGGCT

> A Q L P R A V R V H F M H A V A S G T R W A F L M T V L A G L 20872 CCGCCGCCAGCAGCAGCGCGCGCCGAACTACTAGCGGATTTCCTAGGGTTCCTCGTCGACGCTAGAGCTGAATTCACCGGCGAACTAACA > M S S K I L V I G G G P A G S T A A A L L A R S G L S
21145 GTGACGCTCCTGGAAAAGGAGACGTTCCCGCGATACCACATCGGCGAGTCGATCGCGCTCCTCGTGCCGCACCATCGTCGATTTCGTGGGCGC
> V T L L E K E T F P R Y H I G E S I A S S C R T I V D F V G A
21237 TCTCGACGAGGGTTCGACTCGCGGAGAGAAGAACGGGGTTCCTGCTGCGCTGGGCAACGAGGACTGGGCCATCGACTGGCCAAGA L D E V D S R G Y P O K N G V L L R W G N E D W A I D W A 21329 TCTTCGGTCCGGCGTGCGGCAGGTCGACCGGGACGGCACTTCGACCACGTCCTGCTCAACAACCCCGGCAAGCAGGGCGCCAAGATC
>I F G P G V R S W Q V D R D D F D H V L L N N A G K Q G A K I 21421 ATCCAGGGCGCGCTGTCAAGCGGGTGTTCTTCGACGGTGAGCGGGCCACCGCCGAGTGGTTCGACCCCGAGTCGGGTGAGGTCCGCAC G A A V K R V L F D G E R A T A A E W F D P E S G 21513 CATCGATTTCGACTACGTGGTCGACGCGTCCGGCCGGGCCGGGCTGATCCCGTCCCAGCACTTCAAGCACCGGCGCCCCCACCGAGACGTTCA >K N V A I W G Y W Q G G S L L P N S P S G G I N V I S A P D G 21697 TGGTACTGGGTCATCCGCTGCGCGGCGAGCAGCATCGCTTCGTCGCCACCAGAGCCGCTTCCTGGAGCGGCGCAAGGAGCACGC > W Y W V I P L R G D R Y S I G F V C H Q S R F L E R R K E H A 21789 CTCGCTGGGGGACTGCTGCGCGCACTGGTACAGGAGTCCCCGACCGTGCGGCGCCTGACGGCGACGGGACGTACCAGCCGGGCGTGCGGG >V E Q D F S Y I S D S F C G P G Y F A A G D S A C F L D P L L
21973 TCCACCGGCGTGCACCTCGCCCTCTACAGCGGCATGCTCGCCTCGCCGTCCACCATCCACGGTGACGTCACCGAGGAGGAGG
> S T G V H L A L Y S G M L A S A S I L A T I H G D V T E E E A
22065 GCGGGCGTTCTACGAGTCCCTCTACCGCAACGCTTACCAGCGCTCTCTCGCCGCGCTCTACCAGCAGCAGGCGCAAGAGGGG R A F Y E S L Y R N A Y Q R L F T L V A G V Y Q Q A G K R 22157 CATACTTCGGCCTGGCCGACGCGCTGGTGCACGACGCGCGAACCCGAGTACGAGAAGGTAGACGGGGCCCGCGCCCTTCGCCCAGCTCGTC

>A Y F G L A D A L V H D S G E P E Y E K V D G A R A F A Q L V > A G L A D L D D A A E G R H D S T A A A A P A E 22341 GCTCTTCCTGGCCGCGGAGGGCCCGCCGGATGGCCGACGCGCGCACGCCCGAGCCCCGGTCAGCGAGGCGCCGGGCAAGCTCGACACCC LFLAAEEARRMADARTPSAPVSEAPGKLDS >H D L F D S A T G L Y L V T T P R L G I R R A K P A D T Q A A 22525 GCAGAGCAGTCTGCCTGAGCTTCCACCCTGGTGGCCCCGGCCCGGACCGCCCGGTCCGGGGGGGCTGCTCAACCCTCCCACCACATCCG A E O S A • 22616 GCATCCGGTGCCGGCGCTGAGCAGGGGCAGCGCCACCGACTCCGGCCCGTCACATGGACAAGGTCACCTCTCCCGTGCTGAACAGACGACA 22708 GTTGCTCGCCTCGGCAGCGTGGCCGCGGAGGCACCGTCATGTCACGCTCCTCAGACGGGACGCGCAGGCCGCCAGGCGGCACCGGC > M S R S L R R D A Q A A Q A A P A 22798 GTCGCCGCCAACCCGCACGCCGGGCACGCCGCGCCCCGGTGCCCAGCCGGGTCAGCACGACCACGGTCGCGGTCACCCCGTTCACCGAGCCGA S P A N P H A G H A A P V P S R V S T T T V A V T P F >M P V P P R L T P V S R R D G I D V Y E I P I R P A Q V Q I L 22982 CCCGGCTGCTCACGCCCTACACCTACGCCGGTTCCTTCGTCGGCCCCACCATCCGGGCCCGACGGCCCGGCCCGTGCGGATCACCTA > P G L L T P A Y T Y A G S F V G P T I R A R T G R P V R I T Y 23074 CACCAACGGCTCACCCCACGCCAACGTCACCTCCACGCGCCACCTCCACGCCCACCACGCGACGGTCACCCGATCGACCTCATCCCCGC > T N G L D T H A N V H L H G G H V P A T S D G H P M D L I P 23166 CGGGCGGCTCGAAGGTCTACGACTACCCGAACCTTCAGCGCGGCGCGACGCTCTGGTACCACGACCACACCCCACGCCTACGAGGCCGACCAC
>P G G S K V Y D Y P N L Q R G A T L W Y H D H T H A Y E A D H
23258 GTCTACCGCGGACTGCACGGCTTCTATCTGATCGACGACCAGCACCAGCATCACCTGCGCCTGCCCGCCAGCAAGTACGACGTGCCCGATCAT > V Y R G L H G F Y L I D D P A E H H L R L P A G K Y D V P I M
23350 GCTGCGCACGCCCAGTTCGACGACTCCGGGGCCCTCGTCTTCGGCCACCGGACGACCGGGTCACCATCCTGGCGAACGCCAAGGCCCAGC
> L R N A Q F D D S G A L V P G H P D D R V T I L A N G K A Q
23442 CCTACTTCGAGGTGGCCCCGCGCAGGTACCGGTTCCGCCTGCTCAACGCGGGCGCGTGAAGCACGTCTTCCGGCTCAACCTGGGCGGCGAACCG

>PYFEVAPRRYRFRLLNAALKHVFRLNLGG >V T D P S R V P V T L R A L P P M G T P T V E R T V S M S F D
23810 ATGTCGGCCCGGCCCCCGATCGCGCTCATGGACGGCAAACCGTTCGACCTCTCCGGGTGGACGTCAAGCGGGGCAGCACCGAGAT
> M S A R P P I A L M D G K P F D P L R V D V Q V K R G S T E I
23902 CTGGAACGTGGTCAACGCGGATACCGATCCGTTCCACCACCTGCCACCTGGTGACGTTCCGGGTGCTCGGCCGCGACG
> W N V V N A D T D P F P F D H P F H L H L V T F R V L G R D 23994 GCGGCCCCCCGCGCCGGAGGACGCCGGGCTCAAGGACACCGTCTACGTCTCGCCCAAGGGGTCTGTCAAGATCCAGGTCACCTTCGCCACG SG P P A P E D A G L K D T V Y V S P K G S V K I Q V T F A 24086 CCGTACCTCGGGCAGTACGTCTACCACTGCCACTACCTGGAGCACTCGTCGCTGGGGATGATGGCCCAGCTGGAGGTTGTGCCCTGAGGGC < A A V D V T A P R G P V G C A A R A P L V A P L A A A 25004 GGGTGAAGGCCGCCAGGTCCGCCTTGGTGTGGCCGACGCCGTCGGCGACGACGACGAGAACAGGTCGTCGACGGTACGCAGATCCAGC T F A A L D A K T H G V G D A V A V L L F L D D V T R L D L
 Stope Aggregation of the control of the 25370 CCGGACCCGCAAGACTAGGTGAACCTCTATAGGAATTCGCGTGCCCCCTTCATAGGGTCCGAAAGGGGTAATGGAACCGTCCGGCACCGGA 25462 CGGCTCGTTTTCCTCCCCAATTCCGTCCGACCTGAGCCGTCGCAGGGAAGGCGAGGCCGAGCAGTCGAGCAGTTGATCGGTCGATG 25553 CCGCACGGGCCCGTGCGCCGAAATCGTGGAGATTGCGCAGTGCGTACACCGGATCTGTTCATCGGCGCCCTTCGGCGCCCTTCGTCCCGCCGAC > S T W H Q G P D G W P P H S Y L Q R H L V G G D L L A L E I R 25921 GCAGGGCTGCAACGGGATGTTCAGCGCGTTCGAGCTGGCCGCCACCTCCAGGCGGTACCGGAGCGACCAGCGCCCTGCTGGTCGCCG > Q G C N G M F S A F E L A A S H L Q A V P E R T S A L L V A 26013 CCGACAACTACGGCACCCCGATGGTCGACCGCTGGCGGATGGGCCCCGGCTTCATCGGTGGCGATGCCGGCAGCGCCCTCATCCTCACCAAG 26565 CGGCGTGGTGTGTGCGCGATCGTCCAGGTGATCGAGTCGCCGACGTGGCGGGAGTGACCCGCTGGACACCTGCGGCGGCCCCA
>P G V V V S C A I V Q V I E S P T W R E • 26656 GCCCAGCAAACCGACAGCAGGGGATGATTGTGGAAGCAGAGAAGGACCGGTTGCGTCCGGTGGCGTCCGAGGCGGTCGCCGTGGTGGGGA 26746 TCGGCTGCCGGTTCCCGGGCGACGTCAACTCGCCCGACGAGTTCTGGGACCTGCTCACCGGGGGTCGCAACACCACCGGGACGGTGCCCGAG >I G C R F P G D V N S P D E F W D L L T G G R N T T G T V P E 26838 GAGCGCTGGAGCGCGACCTGGGCTCGGGTCGGCTCCGGTCGGCCACCCGGGCGGCAACTTCCTGGCCGACATCTC

>T L A L A G G V N L I V T P G Q S I T L G S A G A L A P D G R 27390 AGCAAGTCCTTCGACGCCACCGCCGACGGCTACGGTCGTGGCGAGGGGTGCGGCGTCTCGTGCTCAAGCTGCTCTCCGACGCCCAGCGGGA
> S K S F D A T A D G Y G R G E G C G V L V L K L L S D A Q R D
27482 CGGGGACCGGGTGCTGGCCGTGCTGCGGGCAGGCCCAGGCCCAGG > F H S P H M D P L L D R L R A A V D F T A R A P R V P I Y S
28954 CGGCGCTGGCCGACCCGCGGGCCCCGATCACCGCCGACGGCGAGTACTGGGCCGCAACCCGGTCCGGCTCGCCGCAGCGGTG
>T A L A D P R A P I T A D G E Y W A A N L R N P V R L A A A V
29046 GCCGCCGCCGTCTCCGACGGACACCGGGCCTTCATCGAGGTCTCCCCGCACCCGGTGGTGACCCACTCGATCCACGAGACGCTGGCCGGAAG
> A A A V S D G H R A F I E V S P H P V V T H S I H E T L A G S 29138 CCTCGACGACGACGTCTTCGTCGGCGCGCACCCTGCGCGCGACACCCCGGAGGCGCTTCCTGTCCAGCCTGGGGGCCGCGCACTGCC

> L D D E V F V G G T L R R D T P E A Q A F L S S L G A A H C
29230 ACGGGGTCGCGGTCGACTGCGCGGGTCGACCCTGGCGGCACCCGGAGTCACCCTGGCGGCACCCGGAGTCACCTGGCAC 29322 TGGCCGACGCCGGCGCCGCCACGGGCCACGACCCCGCGTCGCACACCCTGCTCGGGGGCACAACGTGGCGGGCAGCGACGACGT > R V W R T A L D D A S R P Y P G S H A L N G V E I V P A A V 29506 TGGTGGAGACCCTCATGGCTGCCGGCGGGGGGGCGGCCGCCGGCTGCTGACCGGCTACCGGCTGATGACCGCCGGGG > V A A I A D H L V E L L S T P Q E .
30517 CGGGCCCTCCTGCTGTCGACGGATCTTCAGGTGGCGGGTCAGCCCGGCCGCTGGTCCACCGCAGGGGTGGCGGCCACTCCAGGTGGCG

<G F I A E L G L A H A G A A F A L V I V V A V A S A A G G D P Q A N A W R M A R R T V P R G V L V A V A L Y A V L Y L L A T
 31252 TCACCACCTGCCGGGGGGTGAGGGTCACCGCCACCGACGACGACGACGACAGCAGGACCAGCGCAGCGTCCTCCAGCGACGCCGCA < V V O G A T L T S V A V S S I L S L L F W A A A D E L 31344 GCCAGGATGATCTGCCCACGTCGCGGTGCAGCAGCGCGCATGGCGAGCGTCTTCGCGATCACCGGCACGGCGATCGCCATCGCCAC

<a href="https://doi.org/10.1001/j.com/red/40/2016/10.20 < G V F L A F V T R E Q G A A L L A A P A L L G A A I G L 32445 CTCCACCAGCAGGTCGAGCAGGGCGAAGTGGCCCGGCTGCATGAACGGCTCCCCACCGGCGAAGTACAGCCGCCGGATGAGGTGGGCGTTCT

< E V L D L L A F H G P Q M F P E G G A F Y L R R I L H A N E RLTQWLEDDDRYADIVASSWAPRKKA 32629 CTGACCGGGTACGCGCACATCACGCACGCAGGTTGCAGGTGTTGCCGAACCGGATGTCGAGGAAGACGGGAAGTCCTCGACGGTGCCGTC S V P Y A C M V C R L N C T N G F R I D L F F P F D E V T G D 32721 GGGCGCGGTGGGGGCGGCCGGTCGGGGTCGGCGATGTCCCCGGAACCGCTGGTTGATCTCCTGCCGGTACGACAGCGCGCGTGGTCCT < P A T R A A L R D P D A I D R F R Q N I E Q R Y S L A G H D !
32813 CGCGGTGGTAGCAGTAGGAGCAGCCTCGCCGGCCAGCATCGCCAGCCGGCCATCTTGGGGCTGTTGAAGGCGTC
< R H Y C Y S C A D V R E G A L M A L R T R R M</pre> 32904 CGCCAGGCCCATCACCCGGCCGGGGTTGTCCGCCGCGTAGCGGGACCGCGGCGAGCAGCCGACGGCGTCGTTGAGCAGGAACTCCGGCT 32996 CCTCCTCCTGCTGCTACAGCTTGTGGGTACATCGAGTCGTCGACGCACCGCCCGTAGACACCGTCGATGGACGCGCAGAGATGGATC 33179 GCCGACCGTCTCGACCCGGGCAGCGGAAAGATCAACCGGGTGCCGCTGGCCAGCATCTCCGCCTCCCGGGCGACGATCTCGTCCCGGAAGT

< G V T E V R P L P F I L R T G S A L M E A E R A V I E D R F H
33271 GCCAGGGCAGGACCAGGTAGTAGTCCGGGCGGGCCGCCGCGACTCCTGCTCGCTGATGATTTCGATGTCCGTGCCGAGGGTGCGCGCCCCCC YDPRAARSEQES 33363 ACCTTGTCCGGATTGCGCTCCGCGGCGTACCGGATGAGCTCGCGGTCTATGCCGCAGAACTGCAGGAGGGTGTTGCCCTTCGTCGACGCGCC 34007 GGAGGGGTCGATGCCCCACCGCTGCGCTCAGGTCAGGTTGCCCAGCAGGGTGCCGTCGTTGCAGCCGATGTCGAGCACCTTGCCGGGCCGCT VRCTTREVV 34464 CTCGTGCCGGTGTGCGGACCAGGCCCCGTCGTCGGGTTTGCTGGCGACCAGCAGGATGTCCAGATAGAAGGGCTGGTCGGGGCCCTCGC

< E H R H A P G P A G D H P K S A V L L I D L Y F P Q D P G E S
34556 TGCGGCCGAGATGCCGGAACGCCCGGTCGAGGTACTCGTCGAGCGCGCGGGGCCGCAGCCGGTCCACCAGGCCGCAGGCCCAGA A R D L Y E D L A R P R L R D V L W L A R L A 34648 CCCACCGGGCCGGACGGCCAGCCGTGCTCGCGGCCGTACCAGCGCAGCAGCAGCAACAGGCCGCGCGGCGCGCGAGGTCAGCTTGACGGT

< V Q T S L V G D F T G D P V P C R G D R D L A H D A P L S 35016 GCCCGGGATGTCGGCGGTCTGCAACTCGGCGGAGCGGAACAGGCCCCGGTATGGCGAGGTGCCCGCGCCGTAGTCGAGCCAGACTCCCGTGG 35473 CCGGCACGTGCGGGCCGAGGTCCAGCTCGACCTCGGCGGCCGGTACTCGAAGTAGATGACCCGGCGCGCTTGCCGGTCACCGCCGGCGCG < P V H P G L D L E V E A P R Y E F Y I V R R R K G T V A P A
35565 GCGTGCAGCATCAGGATGTTGTGCAGCATCACGTCGCCCGGGTTCATCACCGCCGCGCCCCGGTGGTGTCCCACTCGGTGGCGTTCAT</pre> <A H L M L I N H L M V D G P N M V A P V A G T T D W E T A N M 35657 CCGGGTGGTGGTCCCCGGTCGCCGGTCCCAGTAGTTCGACTGCGGGATGCACCAGACGCAGTTGTCCTCCGGGGCCGGGTCAAGGT T D T M T R T R W A E D I L E D A A A O V S R L E D P D L 36117 CCGCGCAGGATCAGCGCGCCTGCCGGCGGAAGGCGGTCAGGTGCTCCGGAAGCAGCCCGGTCTCGTGGATGTGGCACTCGGGGACGGCCTG

<G R L I L A G Q R R F A T L H E P L L G T E H I H C E P V A Q
36209 CTCGGTGCGGACGTCCACAGTCGCGCTCATGGTTCGGTTCCCTTCTGCCAGGCGGACGGTTCGTGCCGGAGCCGCCGGGCCCGG < E T R V D V T A S M < P E T G K Q W A S P E H Q G S G G A P G < G H E F V I V P R W R R L L E V A G R L A L V E G G E <Y P R R L G S Y G P N S V V H V F S D R G T R E A A A A</pre> < V T V G P F D R R L G E A Y S P L A E V A V H R G R P A V R 36852 GCAGGTGACGCAGGATGTCGCCGGCCCCGGATGTCCACGGTGTTGGCGTCCGGTTCGCAGATCTGCTCGATCAGCGCCACGGTGAGC 37035 GGTCCGGCCCCGGAGCCGACCGCGGGGCGJTACGGACCAGGAGTTCCAGCTCCCGCAGCTCGATCTCGGACAGCTCCAGGCCGGCGGCGCGCG G A G S G V A P A T R V L L E L E R L E I E S L E L G A A R TEALPVTPDPWHVSYVDVHDTGLORLSAL 37771 CGCCGGCCGTGGCGATCACGATCTCGTCCCGGTGCGCGGCAACAGGTCGGCCAGCCCGCGGGCAAGGGCCGCCTCGGCGGCCCGCCGCCGC <Y A R S T D F L T V G L D F A R R V A Q V G P E I R R G W Q G 37955 GCCGAGCGCCCAGGTGCCCAGGCCGAGCGCCCGAGACCAGCGGCCCCCGCTCACCGATGCAGCGCTGCCGCACGGCGTCCCCCCCGGTG

```
38689 CGAGCTGCATCTGCATCTCCTCCGGCGTGTTCCACTCCCAGACGTAGTTGAGGTCGCGGTGCTCCGGAGCGGACAGCCCCGGGGGGGCTCC
38873 GAAAACTGGGCGGGGGGGTGTCACGGACGCCGGGCGAACTCGGCGACGTCCTCGGGGGCTCAGCGCGCGGTGTCGGCCAGCCGGGCCAGGAG
39056 CGAGGGCCTCGTCGACCTCCTCGCGGGACGCCTTCAGGCCGGGCAGCGTCTTCGTCGCGCGCCCTGCGGGTCGAGCCGCCGCCGCTAGATCAGG

L A E D V E E R S A K L G P L T K T A G Q P D L R R G Y I L
39148 GCGTCGTAGTTGAGCGCCAGCGACAACTGCGGCACGCCCATGGCGAGCCCGTTCATGTAGCAGTTGGCGCTGCCGTGGTGCACCAGCAAGTC
39424 TTGCCCCACACCACGCAGACCCGCTTGCCCCGGCGCGGACCGAGCAGCAGCGGTCCACGTCCTGGGAGCCGTTGTAGGGCTGGTAGCGGAT
39516 CGGGATCCGCAGCGCCCCATCGGCGGGATCGCCACGTCGGGCGACGGGTCGATGGCGTACCGGATCTGGTGCCGGCTCCACTCGACGC
ARLAWALPVMCMYHGAWNSV
40160 TCATCGCAGCCTCTCTGTGGCCTGCCGAGGGGAGGTTGGGGGTCGCGGCCGGACGGTCAGGAGGTCAGGACCGGCAACTCCCGCGCGGGA
40250 TCGGCGAACTCGATCACGATGGTGCGCCGCCACTGGTCGGAGGGGTTCGGCCCGGAGCCGTGCACCAGCCGTACGTCGTGGACGATGAAGTC
40434 CCTCCAGACAGCCGTTCTCCGGCCCGGCGGTGTCCAGGCAGATGTTGCAGACCGCGTGCGGGGGGACGTTGACCCGGTCCCGGTCC
40526 CACGGCACGCCGGCCGCAGTGGCTCCTTGAGCACCAACGCGAAGGCGGTGGCACGACCGGGGTGCCGAGGACGTCGGCGCCACCGC
<W P V G A A R L P E K L V L A F A T P V V P T G L V D A A V A
40618 GGCGATCTCCGGCCGGTGCAGCAACTCGCCCTGCGGCCAGTCCTGCTTTTCGAGGTTGTGGATCCGGTACAGCACCGGCTCCGCGCCCTCGA

    A I E P R H L L E G Q P W D Q K E L N H I R Y L V P E A G E V
    40710 CCTCGTAGTTCCAGTAGTCGGCGTTGGCCCGGCCGGACCGGCACCGGTCGATCAGGCTTACCGCGCCGGCCTTGAGCTGGGCCAGCACC

            Y D A N A R A P G A F R D I L S V A G A K L Q A L
40893 GAGGCTCACCCCTTCGGTCGGCGGCGTGCCGGTGCGGCGGCGACGGCGATGATGTCGCACACCTCCGGCGACCGGCGCTCGTGCAGC
< L S V G E T P R R T G T R R G V A I I D C V E P S R R E H L
40984 ACCTCCACCTCGGCGAAGCCGGCGTTGTGCAGGCGGGGAACAGGGACTCCCGGTCGAGCCACGCGACGTGAGCCCCCGGGCCTG
        EAFGANHLAAFLSERDLWRVDVSLGRA
I G A A V G A I D K L L T W P E R V H Y L L G A C L V A D F
41260 TCGCCCAGCTCGGTGAAGTCGATCCACCTCCACGTCGGCGACGCGCAGCTCCACCTTGGTGATCCCTTGACCTCCATCACCAGCTCCGCGGG
41810 GGGCCAGTCCCTCGGCCAGCGGGTGTCCGTCCAGTCGCCGAAGACCGATCGGGCCAGCTCCGTCGCGGTGTACGCGGTTCGCACCTCGTCC
```

 < P V G R D V V T Q I G R C H D T V H I W D R V N G G D G Y L P
43187 GCACCCGTCGCCCGTTCAACAGCTCGGTGACGACAGCGGGATCAGCTTCTCCCGGAAACTGGTACGGCCCGTAGTTGTTGCCGCACCGGGTG</pre> V R R G N L L E T V F L P I L K E P F Q Y P G Y N N G C R T 43279 AGGCAGACCGGCAGGCCGTGGGTGCGGGCGTAGGCCAGGGCGATCAGGTCCCGCCGCCGTTCGCCGCGGGAGTTCGGCGCCAG < • V G V R V E S H WYGTIMTSDVRHGRDIMWQVAD 44931 TGGACTGGGCTCTTCGTGAACGTACCGAAGGATCACTCGTGATTTCCCTACTTATGGGCCACCGAGGTGTGATCGGTGGATCTCTATGCGT 45022 CCGCCATTTCCGCAAACGGGGGCCTGGCCGGCCGCCGCCGACCAGTTCGAGCACGGCGGCCAGGCCGTGCACGGTGTCCGGCGGGCTGGCC 45114 GGCAGCAGCAGGGTCTGCAGGCCGGCGTACACCGCCCCACCGTCGGCCAGGGTGTCGCCCACCATCAGCGCCCGCTCCGGGGCGACCTTCAG 45390 GCGGCGTCGCGGTAGGGCAGCCAGCCGTCCGTGCGAACAGCCGGTCGTAGAGCGCCTCGGCGAGGCCGGATGGGGCAGGTCGACGTGGA <A A D R Y P L W G D T R F L R D Y L A E A L G P H P L D V T S
45482 GAGCAGGCCCACGTACGCGCTGCGGTGCGTTTCGGCGAGAGGTCCCGGCGGGGCGTACACCTCGGCGAGCCCCGGCGGGACGCGTGCGGCT 45758 CATCGCTCTCCTCGGTGCAGCCGCGTCAGGGCGGCACCAGCCTGGCAGCGGACCAGCCGGATGTCCATCATGGAGGAATGCGCCGGGTCG 45848 GGCGCGCCCGCCATGGCCGGCCGGTCCGGACGGCATTTTCGGTCACTCTTGCCTTCTAGGCGGATTTCTTCAAAGATGGCTGTCAATTC 46122 GCGTCCGCAACGCGGGCATCCAGTTCGCGTTCATCAAGGCCACCGAGGGTACGAGCTACAAGGACCCCAACTTCAACGCCAACTACGTCAAC
>S V R N A G I Q F A F I K A T E G T S Y K D P N F N A N Y V N
46214 TCCTACAACGCCGGAGTGATCCGGGGGGGTACCACTTCGCCCGGCCGAACATCTCCTCCGGCGCCACCCAGGCCAACTACCTGGCCAACAA

< S H S I A E A A F A A F H F V R E F R N E A F L S D V F D V D

```
SYNAGVIRGAYHFARPNISSGATQANYL
46398 GCACGTCCGGGATGCGTAGCTGGATCCAGGACTTCCTGAACACGTACAAGGCCCGCACGGCCGGTACGCGGTCATCTACACCACCACGAGC
>S T S G M R S W I Q D F L N T Y K A R T G R Y A V I Y T T T S 46490 TGGTGGACCAGTGCACCGGTAGCTGGACCGGGCACCCTGCC
> W W N Q C T G S W T G P W A N H P L W L A R W S S T P G T L P
46582 GGCCGGCGCTTCGGTCTGGACCTCTCGGCAGTACACGGCCTCCGGCAGCGTCTCCGGGATCAGCGGCAACGTCGACCGCAACAACTGGAACG
> A G A S V W S F W Q Y T A S G S V S G I S G N V D R N N W N
46674 GCGACCGCACCCGGCTGATCGCGCAACAACACCTGACCCGAACGCCCGGGTAGGCGGTTGGCGGCAGCGGAACCGATTGCGACCGT
>G D R T R L I A L A N N T •
46765 ACGGTCGGCGGCCGGTCCGGCTGCCGCCACCCCGTCACCCGGCCCGCGTCGCGGGCCCGGCATCCGGTCCTGCCGGCCCGGGGTGGCC
46857 CGCCGTGCCATCCGCCAGGCGGCCACGCCTGCGACGGCGCCGGCCACGACACCGAACAGCGCGAACAGCACCGGAAGACCCGCGAAGACCCGCGCTGACCAG
46949 CAGCAGCACCACGTCGCCGAAGGCGACCATCCACAGTGCCAGCCCGGGTCGAGCGTCCCTGCGTGTCCCCATGTCGCACCTCCTCG
47041 CATCGTCCGGGATCAGATACCCCCTTCGACGCAAGTACATGCGAATCGACATCGCAACGGCGGGAATCGCCTGCTCAGCGGGCCGAGTCGGC
TNQGVSRRVWGKPLPLSFGAPDKHL
L L V A V L A G V L L G R R R R E P K T G G V V T A R G
47868 GAGAGTGCCGGGCCGGAGGCGGCAGCACCGACGTCGGGGGCGGCCTCCTCGCGGGCCGGGCCGGGCCCGGGCCGGGCCGGGCCGG
<S L A P G S A A P L V S T P A A E E R P A V A P E R A P A P P 47960 CGAAACGGGCACGGCCCGGATCGACCCGGGTTCGTCGGCTCGGCTCGGCTCGGCCCGACCGGCCCGGGGGCCCCGTCGGCCCCTT</p>
< G G E P T R P A P P A A P R G T R Q E A P R P A P V V P P R
48420 GGCTCGCGGGGGCGGTTCGGCCCGGGGGGCGGCCGCGCGGAGGGCCG
<P E R P G N P G P R R V G D L L S I T K A R R G A A R R L L R</pre>
EADIREAPDKRLLGNLVPKLPGANR
48604 GCATCGGCTCGGTGGCCAGCGCCGCCAGGLTGGCGATCGCCGACGGTCGGGCGAACGGCGACTTGCCCTCCACCGCCGCGTAGAGCGTCGCG
      M P E T A L A A L T A I A S P R A F P S K G E V A A Y
<G L S W L D A E P G A T G D R A R E P A I Y A P S G L V M G T 48788 CCGCGTCACCTCGGGTCACCCGGGGATGGTCGCCAGGCCGAAATCGGTCAGCACCCCGGCCGTCGGTGCCGAGCAGCACGTTGCCCGGCT
49247 GAGCGCGAAATCAGCCGGTACCGGTCGGCAACGAGTTGGGGAAGCGCGTTCGACATCGGTGGAGACGGTACCCGGCGGCGGCCGCCGCCCAC
49339 ACCGCCGCACGCCACTGTGCGACGAAGGTCAAGTTCGCGACGCTACGCTGAACGGCATGTCTGCCGAAGAGCCGCTGTTCCGGGTGACCC
49799 ATCCGCCAGATGCGCGAGTTCGCCGACAACTCCAGGCGGAGGTCGAGCGGAACTACGCACCGCATCTGTCGTACATAGCGGACGACATGAA
49891 GGCACAGATCCCCAACCCCGCCGACGCCTTCGTCGAACTGGTGCAGTTCCTCAAGGCGCACAACGAGACCCAGCAGGACGCCACCGCCACGATGG
49983 TCTGGTCGTCGGTGGGCCACCGGCCACCGGCCACGGCGGCCGGATCGCCGAGCAGTACGCCGACTCCGACGCCTTCTCCGCCGCT
 50075 CGCGTCTCCGACGTGGAGCGGGCCCTCGCCAACCCGGGCGCGGGGGGGAGGGCCGCCCACGCCGGCTCTGCCCGACCCGACCAGCCCGA
50259 CATGTGGGCGTGCATCCAGGACCACGAGACCGCCGGGCACTGGAAGCAGGTCGCCGGCTGGCGCAAGGTTTGCGACCTCGCGCAGGCCCACC
50351 TGGGCCGCCTCAAGGAGTACCGGCGCGCCTGGCCGAGGCGTGGCCACCGGCGACCAACGCCGCCGCCGCCGCGCGTACCTCGGCGAACTGGAC
50443 GACCTGATCGACAAGGTGCAACGCACCCACGACGCGGCTGCCGCGAACTACGACGCCCTCGCGGGCCGCCACCCGAGCGATCAGCAGCGCCCC
 50535 CACCGAACTCAAGCCGCTACACGACGACGAGTACGTGGAGAAACTCCAGCAGAAGCGGGCGTACGAGGCGACTACCGCCGATCCGAAGGCGCTGA
50627 TGGCAGCCGGCTGCCGGACAAGCCGGTGACTGACGCCGACCTGGAACGCCCAGGCCCGGAACCTGATGTTACGGGTTGACGGC
50719 GAACTCCAACAGGCCCAAGCAATGCTCCGCCAACCCCTCCTCCGCACGGCCCAGGACGACGACAACCCGGATGCTTACGAGAGGACAC
 50903 TTCCTGTCCAGAGCGCACCCCAAACAGGTCCTGTACTCGGGGGGAATCGGGAAACACCACTTCGCCATCGGGCTCCATTTCTCCCGGC
 50995 ACCGGCCACCCAGTGACATCCAACCCGGCAGGAGGTGCTGGCCCTACCCTCCCGCCGAACTTGCCACCTGCCATCGCCCCCACTGGACGAGG
                                                        junction marker
```

```
51087 CAGCACAACCGGCCAGGTAGGAAACCAGCCCCGCTCAACCACGTCCTCACGACCATTCTTGCCAGGCGGCCCGAATCGGCGGGGGCCCGCGA
51179 TGGGATTAATCAACCCCTTGGCAATACTCCTCCGCGAAGGATCAACCCAATTGGAGGGGTGATCGGGGGCGGTGGGGACCTCACCA
51211 ACCGGCGCTGCAGGCACTCGCCCGGGCTCGGGACGTGGTCCCCACTTTAACCTGACCCCTGCGGCGCCTCCGAATGGCGCACCGCAAGGCGT
51363 CGTAAACACCTTAGGTCTGAATGGTGAGCCGTCACGCTCATCCCACCGCGAGCAAGGCGATACTGACGGCCGCTACTGGGATCCCGACCATC
51637 GTCCTGAAACCGATGGTGCCTTGCGAATACGGGCTGACCAGTGGCACCTTAACTACCTGAAGGGGGCGGAGGCTCAGAAGCTGTCACTGGGA

S P E T D G A L R I R A D Q W H L N Y L K A A E A Q K L S L G

51729 GAAGGGGTTGTAGTAGGGGTCCGGGATACTGGCGTTGATCCACACCCCGACCTTCAGCGCAATCTAATCAAAGGGATTGACATCATTCCCGG

> E G V V V A V P D T G V D P H P D L Q R N L I K G I D I I P G

51821 GGGCAATGGAGATGGCCAGAAAGATCGCAACAGTCACGGCACTAGCATGGCTGGACTAATCGCAGCCCACGGACAGGGCCAGAGCGGCCCT

> G N G D G Q K D R N S H G T S M A G L I A A H G Q G Q S G A

51913 TAGGCCATAGCACCCCGAGCCAACCACTCCCATTCCCGCTCGAACAACCTCGTGATGCAGAGCGGCTTGGCTTGCGGTGAGAGAACACCTCGGTGATGCAGAGCGGCTTGGCTTGCGGTTGGAGTATAGAA
>L G I A P R A K I M P I L S S A S N N L G D A D G L A A G I E 52005 TTTGCAATCTCGCATGGGGGGGGATGTCATCAATGTCTCCAGCGGAGGCGGCGCCAGCGTTCGACTCATCAAGGCAATCAGAGAGGCGGTCGC
>G G I D R Q G E H A S V S V V G P E V D L V A P A V D I
52281 AGTTACGACGGGAAGTACTCCAAAGGCACCGGTACGTCCACGGCGCACAGCGATAGTCGCAGGGGCCGCTGCTCTGGTCCGATCGAAGTTTCC
> S Y D G K Y S K G T G T S S A T A I V A G A A A L V R S K F P 52373 CGACCTGCCGCCTCGGAGGGCTCGTCCATCGCCTCACTGCCACCGCCATCGACAAAGGGCCGCCAGGGCACGACCACCACTACGGCTACGGCG
53200 GATGCGGGGCGCAGCGGGGTTCTACACACCGGGCACTGCCTGATCCACGTCATCACGAGACCGCCGAGGCGCCGAGGCGTTCGCCTGACCACCG

> M R G R S G V L H T G H C L I D V I H E T R A E A V A S T T

53292 TGCGTTTCGCTGACATCAGCGACGAGGAGATTGCCGCGTCGCGGTCGCGGCGGCGGTCGCCGGCGGTCGCCATCGACCAT

> V R F A D I S D E E I A A Y V A T G E P L A V A G A F T I D G

53384 ATGGGCGGGGCGTTCCTGGAGGGTGTCGACCGCGCACCCGGCACCGGTGTCGCCTTCCCTACCGTTGCTGCGCCGGCTTCTCGGGGAGCT

> M G G A F L E G V D G D P G T V V 3 L S L P L L R R L L G E L

53476 GGACCTGCGGATCATCGACCTGTGGACGAGGAGGTCGCGCCGGGCGGCCAGGCGGTCGAGGCGGTGGGTACGGTCCAGCCATGACCACGAAGAT
 > D L R I I D L W T K V A P G G Q A V E A V G T V Q P •

53567 CCCTGCCGCTGACCCCGGACTGCGTACGTGCTGGCCCACGGATCGGACCCGGACGAGGTGATCGGGGATCTGATCGAGGAGACCCTC

>S L P L T P E L H A Y V V A H G S D P D E V M R D L I E E T L

53659 GCCGCGCTGCCCGCGGAGGATGCAGGCGCCCCGGAGCAAGCCGGTTCCTGACGTTCCTCACCGGTTGATCGGGGCGCGGGC

> A A L P A E A R M Q V A P E Q A A F L T F L T R L I G A R R A
 53843 AGGAGTACACGGCGTCGCGCGGCGGTACTGGGCGCGGCCGGGCTGGCCGACCAGATCGACCTGCGGATCGGGCCGGCGGGGACACGCTG
 <! K Y A E A Q T P A N L A P K L L P A F L L L Y M V P Q F V
55037 GACGAAGACCCAGACCGGATTGCGGAGCAGGAGTTGCATCTGGCGCTGGGGGACGAGCCAGGTGTCGCGGGGGACTTCATGATCGGACT</pre>
```

V W V P N R L L L Q M Q R Q A V L W T D R A F K M

LGSQDLRRLVQPIATAGEDVTLRLGGGDV 55310 TCCAGCTTGGTGACGTACGGCTCGGTGTCGAGCAGTTGGGCGGCCTGCGGGGTGGCGGCGGCGTCCAGCCCGACGAGCACCTCGCCGGA 55494 CCAGGTAGTGCGTGGTGATGAAGACGGTCATCCCCTCGCCGCGCAGCCGACGGATCTCGTCCCACATGTGGGCGCGACTCTGCGGGTCGAGG CLAGIAGICGIOGICA CARRECTE CONTROLL C L A H Y I S G P Q V N A W K D A G W S L L E R R V D Y G H V 56779 GGCTGCATCCACTTGACCAGGCCGAGAATCATCATGCGAGTGGCAGACACCGGAAAAGCGTATTAGACAAGTTTGACTATCCAAGCATCTG 57328 AGCGGGCCGGCGCGCCCCTGGTTCCCGGTACCTCGGACCCGGTCGGCAGCCCGGACGAGGTGATCGCATTCGCGGTCGACCACGCCTGCCG
>Q R A G A P L V P G T S D P V G S P D E V I A F A V D H G L P 57420 GTCGCCATCAAGGCCGCCTTCGGCGGCGGGGGGGGGCGCGCCTCAAGGTGGCCCGCACGATGGAGAGATCCCGCACCTGTTCGACTCGGCCAC > V A I K A A F G G G R G L K V A R T M E E I P H L F E S A T 57512 CCGGGAGGCGGTCGCGGCCGGGCCGGGCGACTGTTTCGTCGAGCGGTACCTCGACCAGCCCCGGCACGTCGAGGCCCAGGTCCTCGCCG

> R E A V A A F G R G E C F V E R Y L D Q P R H V E A Q V L A

57604 ACCAGCACGCCACGCGACGTCGTCGCCACCCGGGACTGCTCGCTGCAACGCCGCACCAGAAACTCGTCGAGGAGGCCCCCGCGCCCGTTC >D Q H G N V I V V G T R D C S L Q R R H Q K L V E E A P A P F 57696 CTCACCGACGCCCAGCGCCGGCAGATCCACGACAGCGCCAAGCCAATCTGCCGGGAGGCCGGCTACCACGGCGCCGCGCACCGTGGAGTACCT > L T D A Q R R Q I H D S A K A I C R E A G Y H G A G T V E Y L 57788 GGTGGGCACGGCCGGACGACGACCTCCTTCCTTGAGGTCAACACCCGCCTGCAGGTCGAGCACCCGGTCACCGAGGAAACCCGCCGGCATCGACC 58800 TACGCGTACCGGCCGGGGGGTGTTTTCCGCGACCACCGCGAGCGGTGAGGACCGGGGCCGGGAATGATGGCCAGGTGCGGTTCCTACATGGC > > V R F L H G 58891 GCGGTTCCCGCGCACGACCTGACCTACAACGACGTCTTCATGGCGCCGAACCGCTCCGAGGTCGGCTCCGGTTGGACGTCGACCTGGCCAC

> A V P A H D L T Y N D V F M A P N R S E V G S R L D V D L A 58983 CTCCGACGGCACCACCATCCCGCTGGTGGTGGCGAACATGACGGCGGTGGCCGGCGGGGGATGGCCGAGACTGTCGCCGGCGGG > S D G T G T T I P L V V A N M T A V A G R R M A E T V A R R 59075 GCGCACTCGCGGTGATCCCGCAGGACATCCCGATCGAGGTGGTGGCCAACGTGGTCGCCTGGGTCAAGCAACGGCACCTGGTGCACGACACG SGALAVIPQDIPIEVVANVVAWVKQRHLVHDT 59627 AGGCGCGGATGGTCGCCGCGCTGCGGGCGAAGCTTCACCCGGGCGTTCCGGTCGCGGCAACGTGGTCACCGCCGATGGGTA > R D L V E A G A D I V K V G V G P G A M C T T R M M T G V G R > P Q F S A V L D C A A A A R D L G R H V W A D G G V R H P R 59903 ACGTGGCGCTGGCCGCGGCGGCGGCGGCGACGTGATGATCGGTTCCTGGTTCGCCGGCACGTACGAGTCCCCGGGTGACCTGTACACG > D A D G R R Y K E S F G M A S S R A V S A R T A E D S A F D R 60087 GGCCCGCAAGGGGATCTTCGAGGAGGGCATCTCCTCGGCCCGGATGTACCTCGACCCGGATCGCCCGGGCGTCGAGGACCTGATCGACGAGA > A G Y T E G M P L P T S W • 60362 CAAGGGGCCCTTCCTTCGTGCGCGGCTGGGTATCGGCGTGACCGACTGCCGCACGCCGCCGCACACTGAGCCGCCGCCGCCGCTCGAGGGCCC TO THE CONTROL OF THE < H Y G R P G T M
61095 CCTTCGACCGAGGAGAACCCCGATGGCGCCCCTGATCGCTCTCATCGCCGGCTCGGCCCTGACTCGCCGGGCTACTCAACGTCGACG</pre> 61187 CCCTGGCCGGCTGGCACCCCGCCCTGCGGGCCGGGCTCGCGGCCATGTTCGCGCTCACCGGGATCGCCCACTTCACCTCCCGACGCCCGAC 61463 TGACCCTGGCCGGCCGGCGGTGACCCCGCTCGTCCCCGCGGCGCTGCTCCAGTGATCTTCCTCACCGCCGCCGCCGCCGCGATTTCGTTTGGG
>L T L A G R P V T P L V P R A L L Q V I F L T A A A A I S F G > V P L G R L L V T A G H 61737 GCGAGCTGCCGCACCGGGCGGTCGCCGAGGCGTGCTTCATCCGCCCGGCCACCCTAACCGGCATCGTCGACACACTGGAGCGCGAGCGCGACGGCCTC >G E L P H R A V A E A C F I R P A T L T G I V D T L E R D G L 61829 GTCGAGCGGCAACGCGACGTCGACCGGCGGCTCGTCCTGACCCCGCCGGTCGGGAACGGGTCGCGCTCACCAACGT > V E R Q R D D V D R R S V R L V L T P A G R E R V A A L T N V 61921 CATGCAGTCCGGACGACGACCCGGCGAAGGCCGCGTGATCCGGCAGTTCCTGCTCGAGGTCATCGGCAGTG > M Q S G R P M T S V D A D P A K A A V I R Q F L L E V I G S 62013 GAGAGGAACCTCGGGTGACGGCGAGGCCGGAGGCTCCGGCATGCTGATCCGGCTGCTCCGCGCCCACCTGCGCCCGTACCGTC >G E E P R V T A L D A R P E A P A C .
62105 GACCGCTGGCGGGGGTGATGGCGTTGCAGTTCGTCGGCACGATGGCCTCGCTCTACCTGCCGAGCCTCAACGCCGACATCATCGACCAGGG > M S A G M P A E K S M N F G 66424 GACGACCCGGTTCGTGGCGGGTGCCGGCGAGGTGGAGCCGCCCAGTTACTCCAGCTCGTGGAGCATGAGCTGGCGGCGGCCTCGGTGATC <TALEPDTFVNAAVTRLRLPRVAEGLAHWMAI

66883 CCGGCCTGCATGGCGGCGACGCTGGCCAGCGCAACACCCCGGTGCAGTCGCCGGCAGCGTAGATCCCGGGGACGTTGGTGCGGGACACCCC 67067 AGCGCGTGCGACCAGCCGGCCGTCGGCCAGTTCGACCTCCACCCGGTCGCGATGCGCTGGACCCTCTCGGCGCGGGAGTTGTT <LAHSGHVLRGGTGACCCCGGCGGGAGCGCTCGACCCCATGGCGGCGTCGGCGTCGCGCTCGTCGCGGCATCACCCGGTCCCGGCTGGAGA T G A V F T V G A K V L K A H I D A S Q A L A L R K V 67619 TCGTGCACGGCCCGGCGTCGACGGTGACCGCCTCCAGCCCGTCGGAGTGCACCCCGAACTCCTCGGTGTCCCGGTACCCGGTGACCACCTC

<E H V A R A D V T V A E L G D S H V G F E E T D R Y G T V V E < S S A I F T K S P V C D S L V C A G G A G E A E V V</pre> 67893 GCTCACAGTGACTTTCTTCTCCCCGACGCGTCCGACACGCACCGTCGTATTCTCCCCCAGCCGTCCGCCGGGCTATCGTCATCGCCGTGCG 68076 GCTGGCGGCTCACCTTCGCGGGTGAGGGCGCGATCGGCTGGGAGGGCGCGGTCAGCACCATCGTCGAGTCCCCCGGTGATCGGGTGTTCGTG >G W R L T F A G E G A I G W E G A V S T I V E S P G D R 68168 GCGCTCTACGACATCCACCCGTACGACGCCGTCCAGCTCGACGACGACGACGACGCTCCCGGGACGTCCCGCGACGCTCCGCGT > A L Y D I H P Y D A V Q L D E I E G V A S G T Y R K L H V R V 68260 CTCCACCCTCGACGGCGTGACCGCGTGGGTCTACGTCTTCGACGGGTACGAGGGCGGCCTGCCGACGGCGTGGTATCTGTCGGAGATCG > S T L D G D V T A W V Y V F D G Y E G G L P T A W Y L S E I 68352 CCAACGCCGCCGAGAAGGCGGCGCCGACGACGACTACGTCAGCGAGCTGCGGTCCCGCCCCACCGGCACGGCGTCGGCGTAGCGCGTCTC >A N A A E K A G A P D D Y V S E L R S R P T G T A S A • 68443 CCACACTCCCAGTCTGCTCCGCCCGAGACGGGGCCCGCAGGCGCCCCCGCGGGGGTCGTCTGTCACACATCATGGTCGCGCCCGTCACA 68534 CCGCCGTGGCGGGGGGACGGTGCGCTCGTACATGTCGGTCCAGCGCATCTCGCGCAGCCCACCGAGCGGTAGAGCGTCGCCGGGGAGGTC < ATAAPVTREYMDTWRMERLGVSRYLTAPS 68626 GGGTTGGTCAGGTCGACGCCGAGGCCGGCGTGCCGCCGTTCCCTTCGCCGCGTAGACCGTGAAGGCCCGCCACAGCAGCAGCGCGGCGCCGACCCC E G D V E A V F W E D W T R D Y A P L R E R W H D Y G A P E 68902 TAGTCCGGGGTGTCCCGGAACGCCGTGTAGATCCGGTGAACAGGCGCAGGTCCTCCTCGTCGCCGGCGCGCAGCGGCCGGACCGTCAC 69270 GTTGTCCACCGCCGCCGCGGCGAGCCGGCCTGCGGGTCGCTGACCACCACGGCAGGGGTCGAGAAGAGGGGGCCGTCAGGG

< N D V V S W A V L R G Q P D S V L W S D R A P D F F P A T L A
69362 CGGCCTTGACGTCTTCGGCGTCGAAGTCCGGGATCGCGAAGAGTGTCTGCAGGAGGTGCACGAGGAGTCCCAGGAAGTGTCGTCG

< A K V D E A D F D P H G I A F T D A A H V V A L I G P V D D
69454 AGGGTGGGGGGCGCCCCAGTCAGCGGAAGATCACGGGGGGTCCTGGCAGCCACCCCGGTCCCCGCGCCCCTCATTTTCAACCGC

<L T P R R A A W D A P L T V
69545 CCCGCGCCCTGCCCCCGCCGCCGCCGCCCCCGCGCCCCGGGGAGAGGGGACCCTTTCTTCACCCAC 70281 TOGCCOGGCAGCACATCTCCCCGAGCTCCAGCACCAACGGCTTGCTCCGGGCGGCGGCGGCGGCTCATCGGGTTCAGTCT
< A P C R A F M E G L E L V L P K S A A A V Q E P T M < D 70372 GCCCGCGGGGCCTTCCGTGACAGCCCCGTCTTCGGCGTCGACCGCCTCGATCGCCCCTCGATTTCGCCGGTGCGGCGCTGGGCGGCGGT G E T V A G D E A D V A E I A G E I E G T R R Q A A 70464 GACGGCCGTTCCGCCGCCGGCCGGCCAGCTTCCGCCGGCTGAGTTCCTGCTCGGCCTCTGCGCGCGATCGCTCCAGCTCGGCCAGCGACC

```
**R L L R G S R V Q G A V E T D S L A A D L T A E V E G L P L

**TOTAL TREE CONTROL OF THE CONTR
71200 TEGGEGACGGEGGTEGEGGGTEGEGACGAACTTETECEGGEGGCTEGGTGTAGAGECEGEGGAGGAGGAGGTETEGEGGGGGGCGAEGGETE

<E A V A A D R T A V F K E P P E T Y L R R L L S Q P P P V <
72117 GCGATGGGGCGGGAGAGCCTTCGGCGGCGGTGGCGCCCGAGCAGGCGGCACCTGCCGACGGGGGCAAGGCGGAGCAGGGTGCCGATGCCGGG
 72209 GGAACGCGGTCCGGCGGGGGGGGCTCACCGGACCTGCGGGTCGACCAGCGGTCAATCATCTACACCGGAACCATGCGGGTGCGGGTGG
 72943 GCCGCCAGTGCCCGCAGCGGGTCTGCACCATGACCCGGATGCCGACCGCGATGGCCCCCTCGTCGACGTCGAACGAGGCCCGGTGCAGGT
         > P P V P A A R S A P • <A A L A R L A T Q V M V R I G V A I A G E D V D F S A R H L
  74321 CCTGGTGGAAGGGCTCCATAATCTGCGGGACAGCAGGTAGATCGCGGTTGAACGCCGTCATCTGCCCCCCACCTCCTACAACCCGTAACCGA
    74413 TTCGGCGGTCACGAAATCACCGTCGATCCGGGCCGTCGCTACCGAATTGTCGCATTAGTCGTCTCGGTTAACTGCCGCTCGGACAAGTAAC
    74505 CGACCGCACTCGGCAGTCGGACGACTCCCGACGGTGACACGACGGGCCGACCTGCTCCCGCACATGCTGTCCGATCCCAATCACCCGGACGGG
```

74689 CTCATCCGAGACAAGGGGTCAGAACCGGTCGCTGGGGCGGTACGTCCCCCACACCTGGCGCAGGGTGCCACAGACCTCGCCCACCGTGGCC <RARLAEKMPHLVNATGEAAARLEGLAREVAD</pre> 74872 GCTGTCGCGCTCCACTCGCAGCTTGGCCAGCCGCTCGGCCTGGCAGCCTCGATCGTCGGCTCCACCCGCAGCGGCTCGTACGGCTCGTCGG S D R E V R L K A L R E A Q A A E I T P D V R L P E Y P E D A
74964 CGTCGACCGTGAACCGGTTGAGGCCGACCACCACCGCTCGCCCGAGTCGATCTCCTGGGCGATCCGGTACGCGGACTGCTCGATCTCCCGC ZE L Q H Y R A N S E T A T G F G A Y Q R M T W P R S T Y M T TO SECOND CONTROL OF CON < I P F G S E S S R P E S M
76343 CGCACCCCTGTCTTTCCCCGCCGACTCCGACGGTGAACACCTTGGCCACGTTCGCTCCGATTAGGTAAACGTTCCGCCGCGTCGGGTTTCGCA</pre> 76435 TCGGGCGTCGGAACCAGCAAGATAGAGGAGTTGTGTCCCAGCCCCCTCGATTTCCCCCGGTGGCTCTTCTGTGACTCAGATCCCGACGTGGA 76527 GCGGCGGACCAGTCAGCCCACCCACGGACGTGCGGCAGCCGGCACCACCATCGGTGACCGGTACTCGCTCCGGTCCGCGGTGGGCAATGGC > V L G S P H F I S P E R A M G Q E F G P P S D L F S L G V T L 77170 CTACACGGCGGTGGAGGGCCCCCGTTCGACAGGGGCGACCCGATCGAGACCATGCACGCCGTGGTCGAGGACCCGCCGGCCACGCCGC

> C G D G E G K R H G V W R G V V H E G K V Y S F Y L S S T D 78642 CCCGCTTCGCCGAGAGCAAGCCGATCTTCGATCAGATGGTGGCGTCGTTCAAGCTCCGCGGGAGCGACTGAGCCGGGGCCGGGGCCGACGC
>A R F A E S K P I F D Q M V A S F K L R G S D 78733 GACGCCGGCCGGCCGACGCGACGTGGTGAGCCGCCGCCGTGCTATCAAGAGCCATGGCGGCGGACACCACTGACCTCGACGACACG > M A A D T T D L D D 80571 CAGGTTCGTGGTGATATGCGGCCAGGTAAGCCACCGTCCTGCCACTATCCATGTCGTAGAACATGATGCGTGCCACTTGGATGTAGTAGTAG 80663 GTGACCAATGAGGTGAACCCGGGGCCCTGTTTGAGCGTGTACGTAGCGCCGAAGCCACTGGCAGTCACATCGTCACGGCGGGCTTTCCCGTT 80847 GTTCTGCCGGCAAGGGCGAGAGGTCCTGGGCCTCCGTCTTGATCACGATCTCGTAGAGCCGGACATCGGCCGTCTCGGTGTTGTACGCCAGG 80939 ATGACCGGCACCGGCAAGGTGTTTGGGAGCGTGACCGAAATGGACGTGAACCCGGGTCCGGGGCGCTGCGTCGGCCGCTGCCCGGCTTGGGT 81123 TCACCTCGAAGGGCGCTACGGACGACCATCCAGCGCCCCAGTTGTCCTTCTGGCCATCAGGCGGCCACACCTGCTCGCAACCGTCTTCGTTG 81215 TCCACAACCTGCCACACGGCGCTGTTGCCGGTAAGGGCATCGTAGAGCTGGACATACCGCCTGTCGTTCTGCACGACGGCAACGACGACGAGGT 81307 CCACCCCTTCCCGAAATTGCGCTGCCAGCGGGGAGGCCACTCGGGTGGCATCTCGTCGTGGGCATTGACTGGACCGAAGGCGGCCGTGCCGT 81491 TCATAGAGAGCATTCCGGAGTGACATGGGTTTACCCCGTCCTAGGGATCACTGCTTCTCAACAAGATCATCAACGGTGTGCGGCAGGCTGCA 81583 CAATCGGGTGGAAGAAGAGCCAGCACCGGGCCTCACTGGCCATCACAGTATCGTCTGACCTCTTCGACCGGGCCCGAAACGGGCCGAACCGAG 81675 AGCGTTCCGCGCATCTGTCGGTCGGGCGTGTCGAAGGTGCGGCTGTTCGCCACGTGCGCGGGGAGCAGGGTGCCCTCCGCGCCGTCACGGGT 81859 CGGCGGCCCTCCGCGCCGGACCGGGAACGCGCACGCTCACCCCGCGTGATCCCTTGCCCACGAAGGGCTCGTCTGAATAGCATCAGCACAT 81951 CAGGCCCCACCATTACCCGGCGCACCATTGGCATTCGCTTTCTTCGACGTTGCGGCTGCGGTGCCCGGGCGTTGCATTTCCGGCGGAAAGGTGT > A L G A R Q S A E L I D A I R G E C Y E *
82961 CACCCGCAGCGGCACCAACGGCGGGACTGCGTCGAGGTGGCCGACAACCTCACCGGCATCGGCGTCGGGGACAGCAAGGACCCGGGCG 83053 GGCCGGCCTGACCGTCCCGCCCGCCTGGTCCGCCTTCGTCACCGAGGTCAAGGCCAACCGCCTCACCCGCTGACAGCTCCGCAACGAA 83145 CCAACCCCCGGCTACGCCTCAGACGCCAGGCCCCCAGGGCTGCCCCAGGCCTCAGACGCCCTCAGACGTCCCTCAGAC

83237 AGCCCGGGCCGCAGGGGCGCGTGCGCCTCAGACGGCCCGGGCCTCAGGCGCGCTTGCCGAGGGCGGCGTCGACCGCCTTGCCCAGGGCGGT < · A R K G L A A D V A K G L A T 83880 CCGCGGCATGGTCGCCTTGCCGGATTCGTCGTCGTACGCGAGGAGCAGCAATTCCTCGGCCAGCGCAACACCAGTCATGGCCGGAGACGG 83970 TAGCGCCTGAGCGCACCCCGTGCGCCCCCAACTCGCCCACGAGCGCACTCGCCCGGTGAGAGGGGAACCCCGCTATACCGCAGGCGTTAACA 84062 GGGGGCCCTTCCTTGCGATCAGAAGCGGGGCATACCGCCGAACTGGCGGTCGCCGGCGTCGCCGAGGCCCGGCACGATGAACATCCGGTCG CHLAARFNSSDTRADRM 84705 ACGTCCACGATCGCCCAACCTACCGAACGCCCGGCCCCGAGACACGAGGAGCGACGCCGCTCACCCGGCCCGTGCGGCAACCCTGGACGGT 84797 TGGCGTTCGAAGGTAGGGCAGGTGCCCAAGGTCGGGCCCGACCTCGCCGCGTGATCAAGATCACGAGGCGTGCGGTGCGTAGACTTCCGGG 84889 CATGACGGCGACAGCGACGTCGGCCCGGTCGGACCTCTCCGAGCTGGGACGATCCGAGACCGCTTTGCGGAACTTCCTGCACGGCCTGCCG S G R P S G G P D G N A P A G P G V V H L A S V A T A F P S G Q 85348 GGCACCCTGAGGTCAACCCGGGCCGCACCCGGGCCGCAGTCGCGCGGACGAGATCGACATGGTGATCAACCGGGGCGCGCTTCC A P L E V K L A D T R A A V A A G A D E I D M V I N R G A F 85440 TGGCCGGCCGCTACCGCGAGGTCTACGACGAGATCGTGGCACCAAACAGGCGTGGCGACGCCACCTCAAGGTGATCCTGGAAACCGGC T L V M L E A V R D F R A A T G R Q V G V K > R F G A S S L L N D L L M Q R T K L T T G V Y S G P D Y F 85989 GGAGTTCGTCGACCCGGCCGACGCGGCGGCTTCAAGTCGGTCAACCCCGCCTCCGAGGGGGTGCTCGCCGAGATCGCCGAGGCGGCAGCG EFVDPADGGFKSVNPASEEVLAEIAEA 86081 CCGACGTGGACCGGGCGGTCCGCGCCGGCCGGACGGCGTACGAGAAGGTGTGGGGCCCGATGCCGGGCCGGGACCGGGCCAAGTACCTGTTC >A D V D R A V R A A R T A Y E K V W G P M P G R D R A K 86173 CGGATCGCCCGGATCATCCAGGAGCGCTCCCGCGAGCTGGCCGTGCTGGAGTCCCTGGACAACCGCAAACCGATCCGGGAGTCCCGGGACGT > R I A R I I Q E R S R E L A V L E S L D N G K P I R E S R D V 86265 CGACCTGCCGCTGGTCGCCGCGCACCTTCTTCTACTACGCGGGCTGGGCAGACAAGCTGCCGTACGCGGGCTTCGGCCCGAACCCCGGCCGC > V L K P A E T T P L T A L L F A E I C Q Q A E L P A G V V N I 86541 CGTCACCGGCGCGGGCGACACCGGCGCGGGCGCGGGCGCGGGCGCGGGCGCGGGCACACGGCCGGGCGCGGGCGCGGACACGGCCGAGGTCGGCAAGG > V T G A G D T G R A L V E H P G V D K V A F T G S T E V G K 86633 CCATCGCCCGGTCCGCCGCGCAAGAAGGTCACCCTGGAGCTGGGCGCAAGAACATCGTCTTCGACGACGCCCCGGTC >A I A R S V A G T G K K V T L E L G G K A A N I V F D D A P V
86725 GACCAGGCGGTCGAGGGGATCGTCAACGGCATCTTCTTCAACCAGGGGCACGTCTGCTGCGGCGGGTCGCGGCTGCTGGTCCAGGAGTCGGT

> D Q A V E G I V N G I F F N Q G H V C C A G S R L L V Q E S V
86817 CGCCGAGCAGGTGCTGGAGTCGCTTGAACGCGCGAATCGCGGCGTCTGCTGCTGCGCGAACAACACCGACATCGGGGCGATCA

A F O V I F S I K D D M X I I D V G D D I D K N T D I G A I

T F R T P A E A V E K A N N T P Y G L S A G I W T D K G 87185 CCCGGATCCTGTGGATGGCCGACCGGCTGCGCCGGGGTGGTGTGGGCCAACACGTTCAACAAGTTCGACCCGACCTCGCCGTTCGGCGGGSS>S R I L W M A D R L R A G V V W A N T F N K F D P T S P F G G 87277 TACAAGGAGTCGGGTACGCGAGGCGGCCGGCACGGCTGGAGGGGTACCTCGGTGTCTGAGCGGGTCGCGGTACGCAAGACGTAC > V O S A N V S L A S R K 87458 GACGCGGGGACGCCGTGGTCGCCGCCGCGCCGCGCGCGTGAAGGGCTGGGCCGGGGGGCGACCGCGTACAACCGGGGTCAGATCCTCTACCGGGT

> D A R D A V V A A R A A V K G W A G A T A Y N R G Q I L Y R V

87550 CGCCGAGATGCTGGAGGCCGCGCGAGCAGTTCGTCGCGCTCGGCGTGCCGGCGACGAGGTCGACGGGGGGACGACGGTGGGTCTGGT

> A E M L E G R R E Q F V A L G V P A D E V D A A I D R W V W 87642 ACGCGGGCTGGTCCGACAAGCTCCCCCAGGTGTACGGCGGTGCGAACCCTGTCGCCGGGCCGTACTTCAACCTGTCCGCGCCCGAGCCGACG >Y A G W S D K L P Q V Y G G A N P V A G P Y F N L S A P E P T

87734 GGGGTGGTGGCGTGGTGGCCCCGAGGCCCCCGGCGTGCTCGGCCTGGTCAGCGTGATCGCCCCGGCGATCGTCACCGGCAACACGGTGGT

> G V V A V V A P E A P A L L G L V S V I A P A I V T G N T V V

87826 GGTGGGGGGCCTCGCCGACCCAGCCCTGGCCTGGTGACCCTGGCGAGGTGCTGGCCACCTCCGACCTGCCCGGCGGGGTGGTCAACGTCC 88193 CCGGCCGCGGAGGCAGGGGTGGGCGGGTGGGTTGGATCTACTACGAGGGGTAGGATTGCCGCGTGACTCGGTTGGGTGATCTTGAGC 88833 GCGGTCGGGCTCGCGCTCGCGCTCCCCCGATGGGCCTGCCGATGGCGCTCGGCGTGGCCGCTACGACCGGCCGACCGGCAGCGCGTTGCT > A V G L A L G L S A M G L P M A L G V A A Y D R P T G S A L T D L T H G T L P A G L G A V H L G L V G V G F 89017 CGGCGCTGCTCGCCACGACGGTACGCAGCGTGCAGGCGACCGTCCGGGCCCAGCGCAGCACCGGGACCTGCTCGCCCTGGTGGCCCGGGGG >A A L L A T T V R S V Q A T V R A Q R Q H R D L L A L V A R R 89109 GACCCGGAGGTGCCGGGGGCGCTGGTGCTGGACCATCCGAGCGCGGCGGCGTACTGCCGGGCGTGCGGGCCCCGGGTGGTGGTCAGCGC >L A V A V T L V A L P V S L F L S • 89660 CGGACACGTCGCCCGACGTCGCCCGACATGCGGGCCCGGGCCGGGCCGGGCCTGCCGGGCCACCGACATGCGGG 89752 GCGATAGGTAGAGAGCCTACGTGTAGTCTTCCTACGACAAGGGAGCCTACTACCGGAGGGCGGCCATGGATCAACTGCTCCTCGCCCGTC > M D O L L L A R 89842 TCCAGTTCGCCACGACCACCTCGCTGCACTTCCTCTTCGTCGTCACGCTCGGTCTCGTCACCTCGTCTCGTCGTCCACGCCTGGTCACGCTCGGCCTCGGCCTCCAGACGGCCTGG

>L Q F A T T T S L H F L F V V V T L G L V T L L V G L Q T A W
89934 ACGATCACCGGCAATCCCGTCCACGAGCGGCTGACCCGGTTCTGGGGTCAGCTCTACGTGATCAACTACGTGCTCGGCATCGCCACCGGCCT

> T I T G N P V H E R L T R F W G Q L Y V I N Y V L G I A T G L 90026 GCTCATGGAGTTCCAGTTCGGGCTGAACTGGAGCGGCCTGTCGGCGTACGTCGGCAACGTCTTCGGCGCCCCGGTGGCGATCGAGACCCTGG > L M E F Q F G L N W S G L S R Y V G N V F G A P L A I E T L 90118 TCGCGTTCTTCCTGGAGTCCACGTTCCTCGGGATGTGGATCTTCGGCTGGCACCGGCTGGCCCGGGGGCGTGCACCTCGCGCTGCTGTGGGGC >V A F F L E S T F L G M W I F G W H R L R R G V H L A L L W G
90210 GTGGCGCTGACCGCGTACGCCTCGGCGTTCTGGGTCATGGTGGCGAACGCCTGGCTGCAGAACCCGGTCGGCTACGAGGTGCGCGACGGGGT
> V A L T A Y A S A F W V M V A N A W L Q N P V G Y E V R D G V
90302 GGCCCACCTGACCGACTTCGGCGCTTGACCAATCCCACCTTCGGCCTGGCCTTCGGGCACGTGGTCGCCGCCCTGCTCACCGGCG
> A H L T D F G A L L T N P T F G L A F G H V V A A A L L T G 90394 GGATGTGATGGCGGCGGTGAGCGCCTGGCACCTGATCCGGCGCACCCCGGACCACGCGCTGTTCCGCACGTCGCTGCGGATCGGCCTGGTC

>G M L M A A V S A W H L I R R T P D H A L F R T S L R I G L V

90486 ACCGCGGCGGTCTCGATCAGCCTGGTGCAGGGCTTCGGCTTCGCCCAGTTCGGCCGGTCGGGCAGACGCAGCCCACCAAGTTCGGCGGCG

> T A A V S I S L V Q G F G F A Q F G P V G Q T Q P T K F G G G A Q R D A L V A E W T S R F G P G D Y T P P V L A D V G L G 90670 TCATGATCCTGATCGGCCTCCTGGGCTGTCTGTGGCTGCTGCTCCCCCTGCTCTGGCGGGACTGGTTCATCCGGCTGCGCTTCCCGCTC >F M I L I G L L L C C L W L L L P L L W R D W F I R L R F P L 90762 TGGCTGATCCTGCTGCGCTGCCGCTGCCCTTCGTCGCGGTGATCCTCGGCTGGATCGCCCGTGAGGTGGGCCGCCAGCCCTGGGTCGCGTA

V A N W V L F A R Y A A R G A A D P A L G R R P G P A 91038 GAGTCCCGTCCCGTCCCCGTCCTCGGCTGAGGAGGCCCCTGTGGAACTCGCCTGGTACGCCCTGCTCGGGGCTCTTCCTCGCCGGCTAC >A A V A L A A V A G A L S D P V R A A T Q R P L P A V L L P L
91771 GTACTGGTGGCGCGCTGCTGGTGGCCCGGGCGCACCTGCCCGGGGTGGCCTTCGCCGCCACTTCGGCGGCGCTTGGCGCT

> V L V A A L L V A R A A H A R H L P G V A F A A T S A A L A L
91863 GCCGGTGGCGGGGGGGGGGGGGGGGGCGTTGTGGCCCTACGGGGTGTCTCCACCGTCGCACCGACGGCATCACTGAGCGTGACGGCGCGC

> P V A G V G A A L W P Y A L V S T V A P T A S L S V T D A A 91955 CCAGGGGGCGACGCTGACGGTGCTGGGCTGGCGCTACCGCTCCTGCCGGCCCTACTAGGCTTCCAGGCGATGTGCTGGTGGTTGTTCC
>A S G P T L T V L G W L A L P L L P A L L G F Q A M C W W V F 92047 CGGGGACGACCGACGGCACGGCACCGGTGTACTGGTGAGCCGCCGTCCCTTCGACCCACGTCTGCTCCGCCGGGGTCCCCGCGGGCCCGGCG 92138 CGACCTCGCCGTGCTCGCGGTGCTCGGCGGGGCTGACGGCGCTGCTGGTCGTGGGGCAGGCCACCGCGCTGGCCACGGTVCTGGCCGCCGCG 92965 GGTTCCACGCAGCATGGAGGGGCTGGCCGCGGGAGGGCACTGACCCTCTCCGCCGCCGACCCGACCGCACGGCACCGCCACGGCTCGGTCG >R F H A S M E G L A A L D E A L T L S A A D P T A T A T A G S 93057 CGGCCGTCCCGGACGGGCGCGAGATCCGGTTCGAGGGGCGTGACCGTTGCGGGACCGTGGCGTACGGGACGTTGACGGGACGTCACGCTGAC 93425 GCTGGACGAGGTGGCCGCCCTGCCCGACGGGCTCGACACCGTGCTCGGTGAGCGCGGGCACGGCCTGTCCAGCGGCCAGCGGCAGCGGG L D E V V A A L P D G L D T V L G E R G H G L S S G 93517 TCGCCCTGGCCCGGCGTTCCTGCGGGACGCCCGGTGGTGCTGCTGGACGAGCCGACCGCGCGGCTGGACACCGCCAGCGAGGCCGGGGTG >V A L A R A F L R D A P V V L L D E P T A R L D T A S E A G V 93609 CTGGCCGCCACCGGCCGCCGCCGGCCGGCCGGCCGGACCGGACCGGATCCTGCG > L A A T R R L V A G R T A L L V A H R P A L L S D A D R I L 93884 GCCCGGTGGCCGGCGGCAGCGTCCGGGCCGAGCGGGCCGTGCTCCGGCTGCCCGGCCGTACCTGGGCCGGCTGGTCGGCGCGGGGTCT
> A P V A G G S V R A A E R A V L R L A R P Y L G R L V G A G L 94068 GGCTCACCGTGGCGATCGTCGCGGTCCGGGCGCTGGCGATCAGCCGAGGCGTGTTCCGCTACACCGAGCGCCTCGCCGGCCACGATGCCGTG A A F G A T G Y A L D A A A D R A R L A R L E R R L A A 94620 GCTTCGCCGTGGACGCCGCCGGGGCGCTCGTCGCCGGGGTGACCGCCGGCACGGTGGTGGTCACCGCGCTCGCGCGACGGCGTCGCGCGGGGTG >G F A V D A A G A L V A G V T A G T V V V T A L R D G V G G

```
94712 CTGGTCGGGGTGCTGGCCGTCGGCCGCCGTCGAGGTGGCGCTGGCGCTGGGGGCCCCGGCAGCCCAGCTCCGGGC
> A E R G Y L A L T P R P
95723 CGTCGCATGGTGCGCGACGACGACGTACTCGTGAAGGACGGCTGCACGGCTGCACGGCTGCACGGCTCAAGGCCG
WSIGLWEAARTWPPMLVGALVCGAGFFWIG
>I V P L D E D T Y E L D L V V E N L R G G H D T W D S A L L I
97650 GAGCCGGCGAGGTGGCCCGGGACGTCGCGTATGCCCTGCGCCGCGTGTTGGACATGCTCTCCGGCGGCTCCAGCCTCGACGACCTC
> E P A R W P G T S R M P C V C P P C W T C S P P A P A S T T W
TRRCAPRPTAGSGASSAAGG•
97833 TCGGTTGGCGCACCATTGTCGGCAAGATCTCTGCGGTCGTGGACTGGCGCGACTGACACGTTCCAGGGAGCATCAGTCTCTGGCAGAGAAAJ
                        BamHI
                       junction marker
98017 GGTTCGGCGCTGACGTCCGCTGTGGTCGACGACGCAGGTCGGCTGCTCCATGTCTGCGAGATCGGCGACGACCCAGCCGGCTACGCTCAGCT
98109 GGTCGTGCTACTCGTGGAGCGGTCGGGCCGAGCGGGCCGAGCGGCCGATCGCCGCCGACAGCGACCACACGGTCACCTCGCTGCTGAGTG
98201 CCGCCGGGCGTCCACTGGCGATCGCGGACGACTCGGTGGACGACTTCGCCGAGCGGTTCGCCGACGACGACTCGCTGGAGGAGATGCA
```

>LAGYKQVLSAHAALASGRHSAAVALREVLRE

I A R G R E V S V A A D A I A A H L A A D G V A D E G K T V R Q A V A S V R A C D A G C E A L V G A L D A R V T 98844 CCCGGTGCCGGGCCGGCCGCCCGCCGGGGCGAGCCGGTCGCCGAGTTGCCCGGCCCGCCTGCGCGCACTGCGGCCCACAGAGC Q R T I I P P R P K T T G E S A P P P T G G F S A 99488 ACAGCTCCCCGAACAATCCCGTCGCGCGGCGCCCTTGGAGGATCGGCGGAAGCGGCAGATCGACGCGCCGACCCAGGTGGTCCCGCCGGCC >D S S P N N P V A R R P L E D R A K R Q I D A P T Q V V P P A
99580 GAGGGCCGGCTCACCCCGCCCTGGCCCGACGACCTGCCCCAGGAGCCACCGATGCTGCGGCTGGCCGACCG
> E G R V T P P W L A D D L P Q E P P M L R L V E P P P L A D R >G R P A P E P R P E R A F A E H R S P L G Q R V P L E E R P D 99856 ATGGAACATCGGACCGCCCCGCCGCAGCCGTCGCGGTCCGCCGATGGAGCGGCGTACCCCGCCGATCTCCGACGAGGGGGACCGCCACCT > M E H R T A P P Q P S R S A P M E R R T P P I S D E G D G D L 99948 GCTGATCTTCGCCGCCAAGTCGGCCTGGTTCGTCGGGCACGGCGACGACGACGACGACGACGACGCGACACCGGGTGGC > L I F A A A K S A W F V G H G D E S E M D W S S T A D T G W 100040 AGGCCGCCGAGCCGGCCGGCGGGGGGGGCGCCGATACCAAGGCGGGTTGCCCAAGCGGGTGCCGCAGGCCAACCTGGTTCCGGGC >Q A A E Q A A R P A V G A D T K A G L P K R V P Q A N L V P G 100132 TCCCCCCTGCGGAGGGGGGCTCCCCTACGGATAGTCCGCGACGCCGGCCAGCCTCGCCGAGAACACGACCGGCTACTTCCGGGGCTGGCGTCG > S P L R E E R P L R I V R D A A S L A E N T T G Y F R G W R R > G Q E I G G F A V G G R P G R E A A G G W D F T R D T G D R 101417 CTGCGACCAGGCCGTCGCCATGTCGGCGATGAGCACCCTAGCCACGCCGAGCGGGAGCTGCATCCGCGCCGCGATCTCGGCAAGCGAC < A V L G D T A M D A I L V R A V G L P L Q M R A A I 101509 TGCACGCGTCCGTCGCACAGCGCGACGCGATGTACTGGTGCTCTCGGCCCCTCGCCACCGTTGCTACTGGCAGCGGCCCGACCGCGCACCGTCGT <Q V R G D C L A A I Y Q H E R G Q G G N S S A A A R G R V T T 101601 CTCGACGACGCCCCAACGCGCCAGTCCGCCGGGTACGGCCGACGCGCCAGTCGGCTCGT</p> < P V

<G R H Y A S L L G R V G E P T R R Q V S T T P K E V G G P V L
102426 TTGGGCCATCGGCACCCGCTTCGGCAGCCCTTCGGGTGGTCTCCGCCACCGGGACCTCGGTGGCCGCCGAGGCGGCCCGCCAGCCGTCGT</pre> < Q A M P V R K P L G K R T T E A V P V E T A A S A A R W G D 104910 CCGGTCGCCGCCGGCGACGGTCGTCCAGGTCCGACTCGTCCAGGCTGGCCACCGCCTTGAAGCTTTGCAGTGCTGCTGCTGGCCGG

< R D G G A V T Q D H L D S E D L S A V A K F S Q L A Q Q Q G T
105002 TGCCGCTGGCGATGTAGTCGGTGCGAGATGGGGGTCAACTCGCGCTGGATCAGCGCCCGGTGCACCACGACCGGCGGACCGAGAGGTAT <Q D V R I N V R N Y A E Q Y Q A K A Q D G S A G L L L V A S T
105462 GCGTTCGTCCTGGAGGCTGTTGACCAGGTCGCCGAGTAGCCCACCAGATTGGCCAGGTCGCCGGAGCGGTTGGCGTTGTTCAGCGTTTCCA</pre> > M D H P

105919 CACCGCCTCGTCCTGCTCGCCGGCCCTTCGGGCTCCGGAAAGTCGTACATAGCCCAACAAACCGGGCTTCCTGTTCTTTTCTTTGTCGCCGACACTT

> H R L V L L A G P S G S G K S Y I A Q Q T G L P V L C L D D F F

106011 CTACAAGGATGGTGATGACCCTACCGGCCCAAAAACGGGTCTTGTGGACTCGACCCCAGTCGTGGGACGCCGGGGCGCCGTGG

> Y K D G D D P T L P R Q N G L V D W D S P Q S W D A G A A V

106103 AAACGATTGCCCGGCTGGCGCGGGACGGCAAGGCCGAAGTGCCGGTTTATGCGATCGGCGCGGACCGGGGGGCCAACCGGAAATTCCAC

>E T I A R L A R D G K A E V P V Y A I G A D R R V A T R T F E

106195 GTCGCCGGATCGCCACTTTTCGTCGCCGAAGGGATTTTCGCCGCCGGAGATCGTCGACGGAGGGCTGCTCGCCGGGGGCGTA 106287 CGCGCTGCGCCGCGCGCGCACCACCTTTTTCCGGCGGCTCGCCCGCGACCTGGCCGAGCAGCGCAAGGCTCCCGGGATGCTGCTGCGGGC > R V A D L L A G H P H H P • • G L L K G Y A P K I V E D I I 106561 GGCCAGCCGCTCGTCGAACGGGATGAACGCGCTCTTCATCGCGTTGATGGTGAACCATTGGAGCTCCTTCCAGCCGTAGCCGAAGGCCTCCG < A L R E D F P I F A S K M A N I T F W Q L E K W G Y G F A E 106653 CCAGCAGCGCCATCTCCCGGGACATCGAGGTGCCGCTCATCAGCCGGTTGTCGGTGTTCACCGTCACCCGGAACCGCAGATCGCGCAGAAGC TVRFRLDRL < L L A M E R S M S T G S M L R N D T N V BamHI nunction marker < A A L R G L V P P D G P T I D D V I R V G H G L R D A G C W Q 106929 GGATGGCCTGCCAGATCGACGCCCGAACGCCTCGCCGGCGTGAATGGTGAAGTTGTCTCCCGCTGCAGGTACTCGAAGGCGTCC < I A Q W I S P L G F A E G A H I T F H F N E R Q L Y E F A D
107021 AGGTGCCGGGTGGGCGGGAATCCCGCCTCCGCCCCGGCGATGTCGAAGCCCACCACGCCGGCGTGCCGCACCGCCACCTCGGCGAT</pre> < E Q S R A A H R M A T L L T G V R I P H G A D A A L A A G E A
 107205 CGAACCCGGCGACGACCACCTCGTCCAGGGTCAGGTCCCGCTCCAGGTGCTGCTGGGGGGCGAACCGCACCTCGGCGTAGACG
</p> <V G D A A L D L A C E Q A V R R L A P A T Q M V A V T H A F 107389 CTCCAGGTAGCGCTCCAGCGAGCCGGAGTTCGCCGCCGCGACGAACCAGCGGCCGAGCGCTTCCGGGTCGGTGGCAGCTCGTGGCCGA * E L Y R E L S G S N A A A V F W R G L A E P D T T P L E H G V

107481 CCTCGGCGGCCAGCTCGACGATCGTCGCGGCCGCCGCCGCCGTCGAGGTGGTCGTGCAGCACCCCTTGGGGACCTTGACGATGTCCTCG

* E A A L E V I T A P R L G G D L H D H L L A K P V K V I D E

107573 TATGAGATTGCGACCATGCCCAGACCCTAGTAGCGACCGGCGGGGGGGCGGACCGGCTGGGAGGATGTCCAGGTGATGGACCCC <Y S I A V M > M D P > A V G T Y P L V V D A G A G T G R H L A A V L A A L P >R V A G S L A G H F E Q T A E S V L R A R L E L T G R Q V A AVRLGVYRPR• G D R E L E L Y A W D L C Y H L D L L A A A D A P H Q P A 108858 AGGATGCGGGAGCGCACTGCTGGAAGCTCTCCCCGCCGCGATGTGCGGCAGCCGCTCCACCAGCCGCTCGTCGACCGGCAGCGTCGGGTC <L I R S R W Q Q F S E G G A I H P L R E V L R E D V P L</pre> F Q W E P P T V L D V L H S N L L W S M A A Q A P M G F C R 109134 GCCAGGATCACGTGCAGCGCGCGCGCGCCTCGATCTCGACGGTCGGCCGCAGCTCGATCTCGTCGCCCGGCTCCCACACCAGGGGAA 109409 CGGGGGGGGGGGGGTTCAGCCGATCCGCTCGATGACCAGCGGCTGCGGGGTCGGGGCGAGATCCGTACCGCCCGGACCGCCTCG BamHI 109501 GCCAGCGCCGCCGGGATCC

> V A G S P L F V A E G I F A A E I V E E C R R G L L A G A

Figure 11AB

FIGURE 12A

CAG
L
GTC
D
GCC
G
G
GCC
G ATC CGG R CGC R TTC F AAC N GTC CAC GCG R GGC A CGG P DCTG 3CG A > GCG GGC CGA A GAC D GTC V V ATC I GAC CAG GIC GTC CCA CCG GTA CTG L GCG CCA GGT 24 ŋ 3 Ω S 3 ĸ H CGG AAG K CTGACG T GGC AGC GTC $_{
m GGT}$ ggc ACG T CAG GCG CGA S CTG ggc A CGG gcg Д Ц Ø Ы Ö 24 Ø ഷ GCG GGC A G GTT CGG TTC AAG GGC GTC GCC ATG GGC A CGG PGCC CCCGCCAGCCTACGGCCCGGCCGGACGCGGCGGCGGGTGCCGGCGTTGCCCGAGCCCGGTGAGGTTGCTGA A GGG GTG H L Д Ö z × > Ø Ö Д AAG CAG CGC Seg 3TG V A GAC GTC GAG L GGC TTC A GCC GGC Q CGG CTG 4 Ø × Д ĸ Ŋ д 24 CGA GTC ACG T 3TG GAC ggc CAG O C C G ACC T ACC G GGGAAC CTT3AG CGC ഗ Ω U z Д > × D, Ö ᄓ ø > CAC AGC H S GCC CGG GCG 292 299 CCACCA CAC ACC TGC TTC CGC CGC AAC gcc GGC Ø 3 > ĸ Ħ ⋖ z 3 Ö Þ CGC TCC GGC CTC GGG GGTCTG Ø $_{\rm LLC}$ CGC CCC GAC 团 Ø ø ß ഥ Ö α Д Ö വ H ø > SCG ATC I CAC CTC L GAG GTT CCC CGA CGG GTC ATC TTC F ACC CGC ggc ద ы Ħ 2 > z Ö S Ы CAG AAC CAG GGT GIC GTA CGG CGC J CAC IGC AAG TTCGTC CAC acc GAC ى ص ø Ø H > Ω × Д Ö gag 900 GGC CAG GCC GGG GGC ATC AAG CGG R GTC V ATG ATC GCG GGC GAT ĸ Д Σ Ø ø ᆸ Ø × gag TCG ggg TCC R CAC GGT ACC CAG CCA GGC GGT CAG CGC CCG CCC CAG O GCC 3AG CCG GCT ĸ ט Ø Щ > S E ᆸ 3 œ ĸ ß ø CCC GGA CAG GAG ACC GTG TTCCGG CTG3GC Γ CT CCC GAG CGC GTC GAT ø ᆸ Ø > Ľ 24 П Ö Ø Д Ŋ Ö CAG CCC GCC GGG TCC GCG GGG GGGGAC TAT CCG IAC $_{
m GTT}$ Ļ E-ט Д b ы A Ö Ω \succ r Ö ы Z ACC T GAC GAC D ATG M gcc GCG GTC CAG GAT CIC H CGT GTC CTC R GTC GAC Д Д Д Ц ᄓ Н ы ഷ > CGG ACT CTG GIC GTGgcg GAG CGC gag GTC GGG CAC CGG GGT ggc Ø > П > Ö Ø مم ø Д Д Þ ᄀ SSS GGT GGA TCG G S GTC GAG CTC ACG T AAG 3 3 3 CGC A CAG L CTT K GGC CTCCAG CCG TGG GTG d × ٦ EЭ Ц GTG GTC R GCA GCC GCT GCT S GGC GTC V 333 TCC GTA Y CAG L GCG I Ö CCC GAG CCC A TTC F TAC Y ATC I I GCG CAG GGC A GAC CAC GCG GCG Ö J A CGC CAG CAC CGG ATT CAC H TCC S GGC G AAG K K GTG TGC C CGC IGC GTC gcc J Д > ø Ø ď Д Ö CCA TAT ${\tt GTG}$ CCG GGC CAG CAC GGG ACC GAA E CCG CAG CTC3AT CAG L 3 Ö ц Ы > > Ŋ ø 闰 GAC CCG GAT GGC CAT CGT gcg A GAG ATC I GGC G GGC CAG CAG CAG H ы ᆸ > ы ᆸ Ц Ø Σ CTC GGG CAT GAG GTG TTC CAG Q ATG CAG CGC CGG CGC CCG EGAC GGC gcg CGG P Д Δ, Д > Ø α Ø ഷ Ø GAC D CTC L GGG GAC GCC AAC G GTG V ATC I ATC I CGC CAG CGT CTC CCC CTC AGA CGC Ö Н ß Ø ы K ပ GAC GAC CAC CGA CTG GTC CTG 'n GAG E CGC CTG GAC CTC CCG R CTC CGC GTTവ Ω വ z > Ш > > Ø CGG GGC GGC GCC A TCC S T CTG TAG CGG CAC R CTG L GGC A CGG CGC AGC GGT CAC Ö Д Ø ⋖ Ø ы > gce TAC GGC GTA GGC CGA R GGG CTC CCG GAC CCG $^{
m R}$ GTC GGG GGT ATC GGT Ö × Ü 2 Ω 24 Н œ H Д > GGT GAG ggc CAG GGC GTC CCG ICC S CGC A CGG CCA GCC CCC CAG ၁၅၁ GGC ATC I E ы ט Д 3 Ö ᆸ Ø Ø Ø Ü GTC ACC CTG TCC CCG ggg GTA CTCCTC CGG CGC GGC CAC CAG CGG ggc gcc Ω ы ß × Д ഷ Ö > ш Ö ы ĸ Д Ø 858 CCG 1038 CAG 318 CCG 408 GTC 678 CTC 768 GCG 948 GGT 1145 CAC 1235 GAC 1325 GAG 1415 ACC 1505 GGC 1595 GAC 228 CGG 498 CGG 588 GGC Q ^ × ج ы V 1685 ATC < P < A ۷ الا Н Ŋ м У Q v × 四 Ľ Д v 교 凶 ٨ ۸ ٨

1 GTCTTCGGGGAACGCCACGGGAACCTCGTCGAGAACCGGAGCTGACCGTAACCGGTGGGCACCCGCGGGGATCAACCGGTGGTGCTTCCCAGGTTCCGGCGCGGCGGCCCCCG

GCC

Y

Y

Y

CTC

L

L

CTG

L

CCG

P

ACG

A

ATC

CCG

A

ATC

CCG

A

ATC

A

CCG

A

A

CCG

CCCC

CCG

CCCC

CCG

CCG GCC GTC
GTA CCC
Y G
GCG GGC
R A
GGT GGT
T T
T T
CAT CGA
M S
M S
GCC CAC
GCC CAC
GCC CAC
GCC CAC
GCC CAC GGGG P GTC D GCA CTT K CTT L GCG R CGG GAA BGC BGC A BAC CTC CTC BE GCC GC CTC
E
E
GGC
A
A
GGG
P
GGG
P
CGG
V GGT T GGT GCC GAC CCCA V CCCC GG G GCT
S
CCCG
R
GTC
D
GCT
S
GAG
L
CCCC
A
A
GCT
S GTC
D
CAT
M
CTC
E
GTC
D
GGT
T
CCC
G
G
G
N
N rac V CTC
B
GAC
V
V
V
CGC
A
GTC
D
CGG ည် ÀG L GAG CAC CAC CAA L CCG R GTA Y Y AGA S CCC G TTC E GCC G CCG
R
GGG
R
GCG
R
GCG
R
GCG
R
D CAG GTC D GTG H CAG L GAG L GTA Y Y GCAG CCAG L L GTC D GTC D CGA S 3CG R SAC V V CAC CAC D D D D D D D R R R AAG CGT GGT GGG GGG GGG GGG A A A GGG A A A GGG A A A GGG A A A GGG A A GGG A A CTG CTG GGGC
A
A
CGGA
B
A
CGG
P
CTG
CTG
E GGC GAT CCG R CAC CAC
V
CGC
A
GCC
G
CAG
L
L CCG GGGG
P
GTC
D
GAT
I
CCG
R
GAG
L
L
CCG CAT CTC
E
GGG
P
CGA
S
CAG
L
CAG
L 3CA C GCGT T CGT V V GAG L L GTC D D GTC D CGT
T
GGC
A
GGC
A
GTC
D
CGG GAG CCC GGC GCC G GTG H GTC D CGT T CAA > A 2585 GCC > A > D 2135 CAG > Q 2225 ATC GGT GAC 2675 GAG 3238 CAG TCG 2315 CTA 2405 ACC 2495 GCG 2765 TAG 2878 GGC 2968 GCC 3058 CTT 3328 CGA Н H П Ø 2045

CCGCCACCTCCCCGTGTCGCAGGGGACACGCCTGGCGGGTGGTCCCCGGTTGCCCGACCGG R GCG • A GAC CGG V P GCG R ACG $_{\mathrm{TTG}}$ GCG α CCGTTCGGCGAACGTCAGGTCG TCA CGC GCT CAT GTA CAC GAA ACC CAG CAA TCG CCA GAT CGG GCC CCG GCTß ß 3 ĸ CCG CGC -9 9 9 CCG CAG L CGG œ 2 Д CGC ACG GGT GGT CGG GGG Ø CGTTAC CGG CAC CCG CTC > œ CGC CGC gag TTT K CGC Σ ø Ø Ĺ œ CAT GGC A CTG CCC Q GTA Σ ഗ > Д × ŋ GTACCA CGC ACG R GCG R CAG ᄓ > Ø Ø × GTCM GAG L GGT CCG 3AC V CGC A CAT Ö 24 Д Σ Н GGC GTG ATC GAT GGC A GTG H GGA S 3TC D H Ω CGG GAT I ACC GGT CGC A CCG SCC G CCG Д Ľ Н Ж SGC CTC TCC CGC CCA A GTC D CCC 3GC b > Ø 3 ď Ö gcc CAG GTT CGG CAG ACG R CAG 3CG R П Ö z д CTC CCA GGC A GTG CGC A CAG L 3 ы H GGC GCT CAG CGC A GTA ц AGA S GGC GTC ß Д × CGG CAT CTCGAC V GAG L CAG L GAC GCG3GT T TAC V 24 > 山 ACG R GGC GGG L GGGG P CAC CAC CAC വള 3AG L Ø Д CTGGTC CAC Ω Ø GTC ggc CGC CGG CAT CTC 3GG P CAG Д E 3AC Ø Ø Ы Σ CTC agc GCT CGG P GCT SCG R 3AA F CTC 3AG L ß Ø ப CAC CCC CGG
P
GGC
A
CAC
V
CGT
T CCG R GGA > s igc gaa F GAT CTG A GAC V GGC A CGC GGC G CGC ø н GTC CIC CGC S CTC E CGT 200 R 200 A GTC [1] AGA Ø Д H CIC CAC CGC GAT GTC D CAC CAG 3GT T BAC V Ø Ħ > Ч > GAA GCT GTA GCG R CAG TGC ACTC 300 > ᄄ д 闰 ß GTA GAT CCG 3AT I 3GC A CTC SCG CAC 300 æ > 24 × gcg CCG CCA CCG CTC CTC CAC CCG R 24 3 ഷ 闰 ы > TGC CCC CGC $_{
m GGT}$ CGC gcg 3GC ഷ Ø ט Ø Ø ø Н Ø CGC 4052 CAT GCT 4232 CCG GCT GCG ggc GGC 4142 CTG CCA 300 3 ß Ø Ø വ ഷ Ø Ü 3872 GGC 3782 ATC 3962 TAG α ν < A 3598 CCC 3688 GCG Σ ν 999 Д V ۸ ج , L 3508

CAT

GGA

6190 CGG G TCA GGC AGG CCG GGA CTC GAT CCA GTG GAA GTC GGT GGG AGG CAG CGG GGT GAC GTC CTC CAG GGA CAG CCC CGC CTT GCC GCA CAG 6281 CCG TTC CAG GTC CCG CTC CAG GCC CTG CCC GTT GAC CAG CAT CCC CAC GTC GGT CAG GTA CGC CAG CGC GTT GGC CGC CGG
 < GAT GCC CTCGAA GGC GGC GCC ACT CCG GAG GAT TCC CGT CGC GGTCAG CCG CGG 团 GGTഷ ø ø CCC CCG ACG GCTGIC CGTCAC999 GGCບ CCG GGT CGC CAG GTA GCC GTC GCGGGC CAC CAC GCC ACC GTC CCG CAG Jgc GAG ggc CCC GTC CAC GAA GGC CAG CGG CAA GAC CAG GCC ഗ ט CCC GAG CGC CAC CCT ACC AAT GAA GGC gCC CGC GGA GGC ഷ Ö CAG GTC GGT GGTTCG GAG ggg GCG TTG ೮ Д GTC GAC GAC CAC CAG ы > J 6100 GGC CAG GTC GCC GCG TAC GGT TTC GTG CCG GTA AGT CAC ATG CCC TCC CCG GCG CAC CCI ggc CCA CCG GTC CGG GTC GCT CGG CGG Þ Ø Ø GAC CCC ggc AAG GTA CGG CIG CICGTA CGTCGG GTC ĮΉ GGG CGC CGC GAA GCTCAC CAG ď CCC GCA CGG GAC CTCCGC CCC CCC > Ö Ö GCA CGC GTC CTCGTC CAG GTG CAG Ø Ω 团 CTG gaa 6010 GCG CAA CTC CAG CGC GGT TTG CAG GAC CTC GGA GGC GAG CCG CAC CGG GAA ggc CAC GCTД Д Ø GCC CCA GCT CGT GAG GTC CAG GTA GCC GCT GGC CGG CCG AGC AAG Ŋ 団 Д Ü GTG gcc CGC CGC GCG GTC Ö CAC GTA GAG CAG GTA CGC TCC ĸ ഷ GAA GGC GTC CIG೧೧೧ CAT ы CCC < • T A P V G T L
CTC ACC GAC CGG CGG CAC CTT GCG TCC</pre> CAG $_{
m GLT}$ CIC GLICCT CCC GTC 5378 TGACCGGCACCCGG TCA GGT CGC GGG CAC CCC GGT CAG Ö ט CGT TTTGGT CTGCAG GAG gcg CGG Н ĸ ß GCG CAG CAC CAG CAG GCC CAG CAG CGC GAG CAC > GTG GAT CAC ACT 闰 > CAG 335 355 355 GGTGGC 2 Д > IJ CCC A GGC GCG GTGGTC GAG GGC CCC GGT ø GGC GCTGCC GGT GGC Ö Ή > ACC GAC GCC CCC GTG GAC ₽ ᆸ > < • A P R Ą GTG GCC GAG CAG GGC CCA CGC CCG GGC Д Ö 3 Ö GTC CCG CGC GGG ᄓ CCC 5285 CGG CCC CGG CCC Д ט CCG GCG GGC 5920 AAG GCT CGG 回 Д GGC gcc GCA 5560 CAG CAT Ţ GTT ß > Ω υ ט ø 6551 CCG 5830 CGT 6461 CCG 6641 GGG 6731 GGC × ح 6821 GAA 5650 CAT 5740 ACG 5470 AGG

GCTCGTCCGTGTCCGGGA P CAG CAA SAT I SAC V CAA L L GTC CTC
E
GGAA
F
GCC
G
G
GGG
GGG
H
CTC
E
GTC
D
ACG AAC V L GCC GAA CAC (G F V CIC GGGA S CTC E CGG P P GGC A GTC D D 3AA F D 1 GCG CAA L GGTC D CCAG L CTC E GGC A CCG R R GAAA CCG TCA GGGG
P
GCG
R
CAC
V
CGG ECGG Ø CCCC G CGT T GGC A A GGG CAG L CCC G G CC G CCA CCG RCAT M M 3GC A CTC E E CAG L GAC V V GCG CCG GTA Y GGC A CGT T T GCT S GTC D CTC E E GTC AAG SCOT SCOT SCOT SCOT T TCG R CTG Q CAG CAG L L GAC 3GC A CCG R R GGT T GTA CAA L CCG R CAG L L GAA rcg R CAG L GTC D ATG CCG R CGC A GGA S GGT CGC A SAC V CTC E SAC TGA GTC D G G G G G G G TTC GGT CAC V CAG L CGG P CCCA 3GT T 3TA Y 3CT S CCA CTC E E CGC A A GTC D D GTA Y CAT M 3CA C C C 3GC 3GC A A 3GC SAA F SAC V V SCG R R V V SGGG P CGG CGG V V V V V V V CGG P P CGG P CGG P P 300 300 300 300 300 300 CCC G GGC A CGC A TGG I GAA F GGC GGC GGC GGC S GGC S GGTT N CAG GCC D D CCTG O CAC
V
GTC
D
GTC
H
H
CGA CAC V CAG L L CGC AA A A GAT CCG R CCG R R GGC CAG L CCC G G CTG GCG R CGG P CCA W CGG GCG R CAG L L 3GT T T AGA S CAG L J SAC V V STG 3CG R CGG P 3TA Y CCA L L STC D D CGC A A ACC G GTC B GTC D CTG 3GC A CGC A 3TC D GCC G CTC E E CGC A CGG CAG L 3CC G G 3GC A A CGG P A A CGG P P E CGGG
P P GTC
D D D L
L L L L GGCC
G G G G CCGG
R R R R A A CCG A A A A CCC GTC D GGG GGA CCG R GTC D CCA W ATC D GGC A CAC CAC CAC V V 3TC D D CGG AGC A CCG R GGT T GGT 3AG L CTC E CTC CTC CTC E 3AG L L GCT S CAG GGT T GGG R CAG L L CTC E GGT T CAT M TTC
B
GTC
D
CGT
T < CAG GTC D GAC V 7726 CTC < E 8356 GCC < G 8446 GAC < S 8176 GTC $_{
m GGT}$ $_{
m GTT}$ 7636 GAG 7906 CAG 7091 CGG 7276 CGG 7366 GTC 7456 GTG 7546 CGG 7816 CTT DDD 9664 8086 GGA 8536 CCG Q v д v Ц > > ۷ ۲ 7001 7181 8266

12E

FIGURE

FIGURE 12F

 A
 CTC
 CGC
 CAG
 CAG CGT A ACG ATC ATC ATC
I
GAC
D
GTC
V
GTC
V
CGC
R
CGC
R
CGC
R
CGC
R A A G AGC 3CG 3AG 8806 GAT CCG CCA GCT GC < I R W S I 8896 GCC CGA GCG GAC CV < G S R V I 8986 GTC GGT ACC CAG CC < D T G L F 9076 ATT GAA GAT CTT GC < N F I K I 9166 GTT GGC CGT GTC GF < N A T D N 9256 GAA CGC CCC CAG GC < F A G L F CGG 3TG V CTC D D R 999 999 CCCG
R
CTC
E
E
CCCA
W
GCG
R
R
ACC
G
G
T
I
CGT
T 300 TTC
F
TGG
W
GTG
N
O
GAC
D
GCC
D
GCC
B CTC 3CG A > P 10100 GTG GCC > I 9830 GTC GAC 9920 ACG 10190 CTC 9740 ATC F-10010 CCG 10280 GGC ggc > < 8716 CTC 9560 8626 9470 9650

GCG A GTC GCG A CTG CGC R ACC T 10659 CCG GGC C > P G 10749 AAC CTC A > N L

11784 ATGTGCGCGACCCCGACCGGCAGGTAGACCGCCTGGAACTCCTCGCTGTCGAGCACCTGTCCCACTTGCCGAACGTCGGTGAGCCCACCCGCACGTCGAGGACCAGGTCGAGGACCACCC GGCCGTGCGGGCAAGAGAACTTCGCGGTGCCGGGGGTACGGGTGAAGTGCAGCCCCCGCAGCGTCCCCCGGGGGGAGACGCTGTAGCTGGTCTGCCGCACGGGAACAGCGGGTA 12147 GGCGGG 11905